

Towards another dusty summit

The British government seems reconciled to yet another European meeting when it is in a minority of one, but the whole of Europe needs better preparation for the grand schemes now under way.

BRITAIN is likely again to be the odd man out at this weekend's meeting of the European Communities (EC) in Luxembourg, but for no good reason. It is a curious, but explicable, situation, turning on the exact meaning of the word 'federal'. A year ago, the then British government under Mrs Margaret Thatcher was forced (some say "ambushed") into participation in two conferences (on political and monetary union). The objective is to amend and strengthen the Treaty of Rome, upon which the continuation of the European enterprise depends. Now there has appeared (not for general circulation) the draft of a document whose preamble contains the word the British distrust. The fear has taken root that the Prime Minister, Mr John Major, cannot assent to such a piece of paper without further alienating those among his own supporters who nurse the deepest suspicions of the mainland, but cannot refrain from signing it without risking further derision and contumely from the other eleven members.

The weekend's row, if it materializes, will be symbolic rather than real. There is nothing intrinsic to the use of the word federal that need so offend Britain's Eurosceptics. The dictionary meanings of the word range from one that covers the EC as it is at present (all of whose members have ratified the Treaty of Rome) to one that described the United States of America two centuries ago — states compacted to act externally in concert which are nevertheless internally independent — but which is an insufficient description now. In the real world, Major could agree to the offending piece of paper while emphasizing both the conditions that must be satisfied and the time that must pass before a true federation will be possible. If reason were all that mattered, he could prudently calculate that the chauvinistic inclinations of many other member states will provide ample breathing-space.

Sadly, the issue that has now been sharpened is not entirely a matter of reason. Britain has never been a full-throated member of the EC. It stood apart from the negotiation of the Treaty of Rome in the 1950s, responded to the rejection of its membership application in 1963 by setting out to create the European Free Trade Area, but abandoned that when, in 1970, it seemed as if membership would be preferable and possible. For much of that early period, British ambivalence had its roots in hankerings after stronger Commonwealth and/or Transatlantic relationships — hankerings themselves undermined by other ambivalences. Since the early 1970s, there has never been a British government so sure of its own party's European convictions that it could warmly advocate the benefits and the opportunities of being a part of Europe

proper, not just an off-shore island. The result is that Britain has been obdurately, even aggressively, lukewarm about European membership. Major cannot complain that the Luxembourg meeting promises to be yet another occasion when the other members seek some more positive sign.

The trouble is mostly political. The Conservative Party's divisions on European policy have been sharpened by the departure of Mrs Thatcher after a row on precisely this issue after Sir Geoffrey Howe's resignation last year. Major has been hoping for peace and quiet while the negotiations are under way, promising that it will by then be for the House of Commons to decide. But that is clearly a high-risk strategy. Would the Eurosceptics toe the line any more willingly then, closer to the deadline for a general election, when many of them would have even less to lose by opposition?

The pity is that the party division probably does not represent a similar difference of opinion within Britain as a whole. And politicians are probably as bewildered as their electors by the heady European concepts now in the air — that of a single currency (on which even Germany has now cooled) and of the "economic convergence" that is its pre-condition, for example. And what exactly is meant by a federal Europe when there is no agreement yet on the technicalities of defence, let alone on broad questions of foreign policy? What Britain and the rest of Europe need is plain speaking about all these issues. British politicians may have an immediate interest in keeping quiet on them, but those elsewhere have not been all that open on the same issues. They should more openly speak what is in their minds.

Early bird born late

A pre-*Archaeopteryx* animal with wings will have to make its way with difficulty into the world.

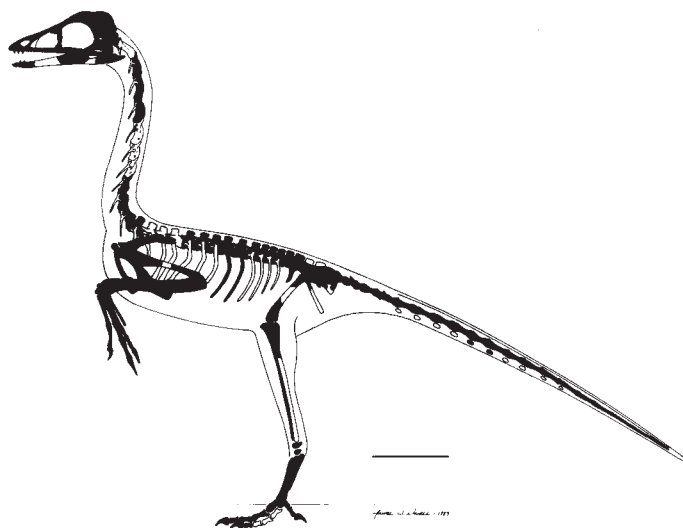
THE timing of the discovery of *Archaeopteryx* in 1861 was impeccable, just when Darwin's *Origin* needed support in the eyes of critics. But that is not how it seemed to Richard Owen, the anatomist who remained in ironic opposition to Darwin's views despite his formal description of *Archaeopteryx* ("ancient wing") in the *Philosophical Transactions of the Royal Society*. From the outset, *Archaeopteryx* was not just the missing link between reptiles and birds, but a source of controversy, the subject of allegations of Victorian scientific fraud and the very emblem of evolutionary transforma-



tion. Now it is ingrained in the public consciousness as the epitome of a fossil, an exquisite testament both to evolution and to a plucky little bird launching itself unsteadily into a hostile world, endowed with a preternatural sense of responsibility for its innumerable descendants. *Archaeopteryx* has become a kind of Jurassic Otto Lilienthal.

Only a direct challenge of its status as primordial bird could unseat it from its perch, but nothing less is promised by an article to appear this Saturday (29 June). It describes a bird far more birdlike than *Archaeopteryx* that was alive and flying 75 million years earlier. Fittingly, the article appears in *Philosophical Transactions* (332B, 277–342), and will cause similar consternation. The evidence for the new early bird consists of two partial skeletons from the 225-million-year-old Dockum formation of Texas which, if they are validated, would make *Archaeopteryx* merely a throwback in a world otherwise presumably teeming with advanced birds, as much a relic of a vanished age as is the platypus today among mammals.

So palaeontologists will be as sceptical of what Sankar Chatterjee of Texas Tech calls *Protoavis texensis* as was Owen of his own creation. Even by the leisurely standards of the field, *Protoavis* has been a long time hatching. The fossils were found as long ago as 1983, and their progress into the literature has been marked by a lone and tantalizing abstract, itself now four years old. Even the monographic treatment now given it will fail to satisfy: it discusses only the cranial anatomy in any detail. A discussion of the postcranial anatomy — and the all-important flight capability of *Protoavis*



Protoavis texensis, n. sp. — composite skeletal restoration representing the size and proportion of a large individual. Scale bar, 5 cm. (Reproduced with the permission of The Royal Society.)

which, incidentally, does not arrive clothed in feathers — is promised for later. The presence of a few more bird bones from another site in Texas are hinted at even more elliptically. Yet again, these are promised for a separate paper. Palaeontologists who have known about *Protoavis* for years, but have had to make do with rumour and guesswork — and even restorations in popular books — will want to know much more before they will believe it is the real thing. Perhaps Owen, after all, will have had the last laugh.

Back to old Berlin

The choice of Berlin as the new capital of Germany is good for Germany and for the rest of us.

LAST week's decision by the German Bundestag that the German capital should be moved from Bonn to Berlin may have surprised the pundits and dismayed most of those who work in Bonn, but it is both sensible and imaginative. The most immediate benefit will be that accruing to the eastern states of Germany, for which it is economically, culturally and politically important that something more should be made of the city of Berlin than a lasting monument to the period during which Germany was divided.

And the geographers' complaint that Berlin is "too far East" merely means that it is inconveniently placed for those who already have well-beaten paths of travel from which they will deviate only with difficulty. It is, in any case, as fair to argue that a capital not far from the border with Poland and with Czechoslovakia will help not just Germany, but the rest of Europe, look eastwards, where there is a great deal of interest.

The canard that making Berlin the capital of Germany will serve only to restore German militarism and a sense of power is a more interesting but even less worthy objection. The desultory, but absorbing, items published in our Correspondence columns in the past few months about the causes of the Third Reich and its aberrations have uncovered little new, but may have reminded many readers that Germany's traditions are as liberal as most would reasonably ask. And even in the twentieth century, Berlin is more readily recalled for Berholt Brecht as for Adolf Hitler, who went north with something of the diffidence that Bonn's civil servants will now move east. But that will take a long time. Could science ease the path by using the designation of Berlin as a way of creating the membership disciplinary societies most conspicuous in Germany by their absence? □

Jobs wanted?

Nature is this week launching a new service for readers — from this issue, we shall include a 'Positions Wanted' section in the Classified Advertisements section.

It is not merely that economic ups and downs (mostly in the West) and political upheavals (in Eastern Europe, the Soviet Union and the Middle East) have left some professional people without jobs, but that recent research among *Nature's* 500,000 readers suggests that 60 per cent of them would be willing to move to positions in other countries.

The Positions Wanted column is meant to meet international needs, but should also cater for those who would prefer to stay at home. To launch the service, we shall publish two advertisements for the price of one. Full details may be found on Classified page 18. □

University of Tokyo set for major reform

- Graduate schools to get more influence
- Other universities may follow suit

Tokyo

AFTER years of discussion, the University of Tokyo, Japan's leading national university, seems set to undergo radical reform designed to strengthen graduate level research in the university. If successful, the changes at Tokyo are expected to trigger similar reforms throughout the whole of Japan's national university research system.

Over the past few months, Tokyo University officials have been drafting detailed plans for the reform, which the Ministry of Education, Science and Culture will submit to the Ministry of Finance in August in the education ministry's budget proposal for next fiscal year. Next year's changes — if approved — will be made in the university's faculty of science, various laboratories and research institutes affiliated with the university, the faculty of engineering, and the college of arts and sciences, which together embrace thousands of graduate students. They follow similar changes in the faculty of law this year, which only affected a very small number of graduate students.

Under the present system, all the power (and most of the budget) of Japanese national universities is concentrated in undergraduate education. The graduate schools are treated just like extra appendages, and in the present environment of government fiscal restraint they are consequently starved of funds. This situation is particularly odd for science graduate students at Tokyo University because their number (1,172), including those at affiliated research institutes, is nearly twice that of science undergraduates (660). The university's faculty of science has been pushing for reform on and off for thirty years, but only recently has its voice been heard (*Nature* 338, 99; 1989).

The proposed reforms are designed to turn the present system on its head by concentrating power and funds in the graduate schools. Each member of faculty will cease to hold a position in the undergraduate school and will instead be assigned a position in one of the university's graduate schools, where their prime responsibility will be to supervise graduate level research. Most faculty members will continue to teach undergraduates, but some will be engaged purely in research and research supervision (a situation that already exists in the university's affiliated research institutes).

A new graduate school of sciences will be created in Tokyo University by drawing several hundred faculty members from the faculty of science, college of arts and sciences, and various laboratories and research

institutes affiliated with the university. A new graduate school of mathematical sciences will also be formed with staff from various faculties. The faculty of engineering will have its own reformed graduate school.

Another key change in the graduate school of sciences will be the introduction of large multidisciplinary research groups. In general, the Japanese university research system is based on small independent research groups (or *koza*) typically comprised of one professor, one associate professor, two research assistants, and a small troop of graduate students.

Young researchers usually become closely tied to a particular *koza* in their undergraduate years and tend to stay with that *koza* throughout their research careers. This creates a very rigid in-bred research system where professors reign supreme over tiny research groups that have very narrow research interests. There is very little interaction and exchange of researchers between *koza*, even in closely related fields. The system is perpetuated by the ministry of education, which hands out tiny fixed budgets in piecemeal lots to each and every *koza*, regardless of performance.

The university's plan calls for five multidisciplinary research groups in the new graduate school of sciences, covering computational science, earth and planetary science, artificial materials, superparallel computer systems, and molecular biology.

Each group will have approximately fifteen to twenty faculty members but only

about four will be permanently assigned to a group. Professor Ikuo Kushiro, dean of the science faculty, says he and other reformers hope that the multidisciplinary *koza* will offer greater choice and flexibility for young researchers joining the system.

The reform should bring in extra research funds for the *koza*. When the faculty of law underwent reform this year, *koza* funds for faculty members engaged in teaching and research were increased by about 25 per cent, and other faculties are hoping for a similar boost. But the increase requires the support of the ministry of education and the approval of the ministry of finance and a decision is not expected until the end of this year.

The ultimate aim of the reform, says Kushiro, is to turn Tokyo University into an international centre of excellence, such as the US Massachusetts Institute of Technology or the University of Cambridge in the United Kingdom, by making it an attractive research environment for non-Japanese.

"This is a wonderful idea, in principle," says Robert Geller of the university's geophysics department, who is one of only a handful of non-Japanese faculties to obtain tenure in the university. But, although he is sure that the reformers are sincere in what they say, he wonders whether the majority of faculty members are committed to opening up the university to non-Japanese faculties.

Many observers both within and outside the university suspect that the prime motivation for reform is the desire to get more money. But if money were everything, Geller says, "Saudi Arabia would have the best universities in the world".

Whatever the case, the changes at Tokyo University are being watched very carefully by all the other major national universities. If the reform succeeds, others are bound to follow suit with their own proposals for change.

David Swinbanks

The trials of young researchers

Tokyo

In an unusual move to back up calls for reform of Japan's national universities, deans of the science faculties of ten major national universities released a report in March that details the problems confronting Japan's young university researchers.

The report charges that the creative talents of young university researchers are stifled by the lack of technicians, the lack of laboratory space, the lack of research funds, and by a rigid research system that prevents them from freely pursuing the research of their choice. This may not be new, but it is unprecedented for such a large group of powerful senior academics to admit the problems so publicly and so frankly. The committee consisted of science deans from the universities of Tokyo, Kyoto, Nagoya, Osaka, Tsukuba, Hokkaido,

Hiroshima, Kyushu, Tohoku, and the Tokyo Institute of Technology.

The deans are blunt in recommending solutions. Graduate students should be allowed to apply for their own research grants (at present they cannot), they should be provided fellowships (there are very few fellowships currently available at the graduate level), and the ministry of education should introduce a formal system of teaching assistantships in line with recent moves in this direction by some of the universities.

The deans also call for re-examination of the *koza* system, the small research groups that form the very framework of the university research system (see above). And they suggest that the lowly research assistants on these teams should be given greater freedom by, for example, allowing them to teach graduate courses.

DS

Canada gets cold feet

Washington & London

IN a surprise move, the Canadian National Research Council (NRC) has announced that it will pull out of a major joint project with the United States and Britain to build two 8-metre optical/infrared telescopes.

The two telescopes, one in the Northern Hemisphere at Mauna Kea in Hawaii, the other to view the southern sky, probably at Cerro Pachon in Chile, will cost \$176 million to build, and are scheduled to be completed by 1998. The plan was for half of this money to come from the US National Science Foundation, with the Canadian NRC and the British Science and Engineering Research Council splitting the rest of the cost equally. Canada's withdrawal should not delay construction of the Hawaiian telescope, which should begin next year, but if the second telescope is still to be built, another partner must be found to replace the Canadians.

Sidney Wolff, director of the US National Optical Astronomy Observatories, says she

is surprised by the timing of the decision and by the lack of prior consultation with project leaders in the United States. She had expected the NRC to defer its decision until two joint Canadian Natural Sciences and Engineering Research Council/NRC committees delivered their final verdicts on the telescope project. Instead, the NRC looked at preliminary versions of the committees' reports and decided that the project was too expensive.

In a letter sent to Canadian astronomers, Don Morton, director-general of the NRC's Herzberg Institute of Astrophysics, explains that the Council had to assume that its astronomy budget would not be increased. (The Canadian government is trying to control public spending in order to reduce the national budget deficit.) Given this, the letter says that even small cost overruns on the 8-metre telescopes would threaten Canada's other astronomy projects.

Canadian astronomers are lobbying hard for the project to be reinstated. Project scientist Gordon Walker, from the University of British Columbia, fears that if Canada withdraws from the 8-metre telescope project, its astronomers' access to US telescopes in the future will be restricted.

Morton says that the NRC's decision is not yet final. But a reversal seems unlikely, unless the Natural Sciences and Engineering Research Council or a foreign funding agency agrees to share the NRC's costs. Wolff is not optimistic about a Canadian change of heart, and is already looking for other partners to join the project. Morton's letter is "the firmest 'no' I have ever seen from a government agency," she says.

Wolff says she is disappointed by the NRC's move, but does not expect the project to be delayed as a result. The US and British contributions are enough to pay for the Mauna Kea telescope, and with no spending due on the Southern Hemisphere telescope for another two years, there is time to find a replacement for the Canadians, says Wolff.

The Canadian withdrawal is not the first funding problem to hit the 8-metre telescope project. Last year, the British contribution was delayed for two years, as part of the drive to prevent a £40 million budget shortfall at the Science and Engineering Research Council. Until the British can afford to contribute, the National Science Foundation is funding the project alone.

Many US astronomers would have preferred the project to have been funded in full by the National Science Foundation. In March, a National Academy of Sciences committee under John Bahcall, from the Institute of Advanced Study at Princeton, made a US (rather than a collaborative) optical/infrared 8-metre telescope at Mauna Kea its highest priority for funding in ground-based astronomy for the 1990s.

Peter Aldhous & David Lindley

Japanese agency's faith in HOPE

Tokyo

JAPAN'S National Space Development Agency last week submitted an ambitious proposal for an unmanned space shuttle to service the Japanese module for the planned US Space Station. The proposed shuttle is twice the size of the one originally designed by the space agency and will be able to carry three times as much cargo into space.

The proposal came shortly after the US House of Representatives renewed its commitment to the space station by voting to spend \$1,900 million in fiscal 1992 for its development (see *Nature* 351, 507; 13 June 1991). If the Japanese government accepts the space agency's proposal for an enlarged shuttle, it will add considerably to Japan's already large financial commitment to the space station.

For several years, the space agency has been drawing up plans for an unmanned cargo shuttle to be called the "H-2 Orbital Plane", or HOPE, after Japan's next-generation H-2 rocket that will carry the shuttle aloft. Last week's proposal, however, is the first time the agency has requested a budget to begin development.

The space agency has asked for ¥20,000 million (\$140 million) for the first three-year phase of research and development, which would start next fiscal year. Agency officials are unwilling to provide total cost estimates for HOPE at this stage, but it will certainly add significantly to the ¥300,000 million (\$2,100 million) that Japan is already committed to spend on the Japanese module for the station. The shuttle is expected to be completed by 2000.

Earlier plans called for a shuttle 7 metres long weighing 10 tons. But in the proposal submitted last week to the Science and Technology Agency, the space agency requests a shuttle 18 metres in length weighing 20 tons. The agency decided to increase the size of HOPE, says Mutsuhiko Masuda of the agency's public relations division, because a 10-ton version would only be able to carry about one ton of cargo to the space station, and this would limit activity in the Japanese module. The 20-ton design will be able to carry 3 tons to the station and can carry 5 tons back to Earth.

Science and Technology Agency officials are not prepared to comment on the likely success of the agency's proposal. The proposal will first be examined by the Space Activities Commission, Japan's top space policy-making body, before the science agency decides whether to include it in its 1992 budget request, which will be submitted at the end of August.

David Swinbanks

ENVIRONMENT

US blocks Antarctic accord

Washington

THE US delegation to a meeting of the Antarctic Treaty nations in Madrid last week blocked plans to sign a protocol that proposed a ban on mining in the continent for 50 years or more. As expected, the United States was alone among the 26 consultative parties to the Treaty in opposing the protocol (see *Nature* 351, 88; 9 May 1991). But because the Treaty requires consensus on any new protocols, the US position ended hopes of holding a symbolic signing ceremony on 23 June, the thirtieth anniversary of the enactment of the Antarctic Treaty, under which conflicting territorial claims to the continent were put aside.

State Department officials say that the United States could not sign a document that would have allowed a minority of Treaty nations to block a majority decision to relax a mining ban after 50 years. Under the draft discussed in Madrid, a decision to allow mining would require the agreement of all the present consultative parties. The US delegation instead put forward an alternative draft. This would allow a nation to withdraw from the mining ban, if a majority supported an amendment but a minority of nations were blocking the decision.

As a compromise, the other Antarctic Treaty nations changed this walkout clause to require a three-quarters majority in favour of an amendment. But the US delegation refused to sign even this version, saying that more time was needed to consider its implications. Peter Aldhous

PARTICLE PHYSICS

SSC gets new detector

Washington

SEVERAL dozen institutions, most of them in the United States, have agreed to collaborate on building a new major detector for the Superconducting Super Collider (SSC). The detector, if it is approved by the SSC Laboratory, will take the place of the L* (pronounced 'L-star') detector, which was cancelled last month because of problems with its funding and its management (see *Nature* 351, 258; 23 May 1991).

The physicists building the \$8,200 million SSC desire to have two separate major detectors that will observe the accelerator's particle collisions in complementary ways. One detector, being built by a group called the Solenoid Detector Collaboration, has already received approval from the SSC Laboratory; it will be a general-purpose detector designed to track large numbers of particles at once.

The cancellation of the L*, which was intended to track fewer particles with greater accuracy, left the SSC with only one major detector in the works. In response, SSC Laboratory director Roy Schwitters invited all interested physicists to a Dallas meeting 11–13 June to discuss the possibilities for building a successor to the L*.

The meeting was as successful as Schwitters had hoped. By 18 June, a collaboration had formed that included most of the scientists that attended the meeting. The interim leaders of the group are Barry Barish of Caltech and Bill Willis of Columbia University.

The collaboration is still growing, Barish says, but already it has physicists from 35 to 40 institutions. They include "a good fraction of the high-energy physics community not involved in the other detector" and "almost everyone from L*", he says. Few of the institutions are from outside the United States, however, and Barish attributes this partly to the problems surrounding the L* collaboration. Germany, Switzerland and the Soviet Union resigned from that collaboration, complaining that they were not being given enough respect.

To date, the major foreign partners in the new collaboration are China and Rumania, and Barish says the Soviet Union, Korea and Taiwan have also expressed interest. As time goes on, he says he hopes some European partners will materialize.

The group is proposing to build a detector that will do many of the same things that the L* would have done. That makes sense, Barish notes, because the L* was not cancelled for scientific reasons. In particular, the new detector will be designed to accurately identify and measure electrons and muons produced by collisions in the SSC. It will not be able to follow the tracks of the particles as accurately as the general-purpose detector, but it will provide more accurate information on the particles' properties.

Barish says the collaboration will be care-

ful to avoid the cost problems that plagued the L*. The SSC Laboratory wants the two major detectors to cost no more than \$500 million each, but by the time of its cancellation, the price tag on the L* had ballooned to more than \$750 million. (The cost of the general-purpose detector is now estimated at more than \$700 million, and its designers are now trying to pare it back.) Barish says his group is already discussing ways to keep the cost down without hurting performance, such as using a superconducting magnet, instead of non-superconducting as was planned for the L*, and not installing steel shielding around the magnet.

When the L* was axed, Barish — who had played a major role in the collaboration — quickly contacted many of the participants to ask them to try again. Within days of the cancellation, Barish held an informal meeting at Caltech with most of the principals plus others who had not been a part of the L*. "I felt that if there was any time lost at all, [the collaboration] would fall apart," he says. By the time of the meeting at the SSC Laboratory, the replacement was already mostly in place.

There remain a series of steps necessary before formal acceptance of the detector by the SSC Laboratory, but final approval seems a safe bet — this is the only good chance left to build a second detector for the SSC before its opening, now scheduled for 1999. Although Barish's group is starting a year late, he hopes to make up at least half of that and to be no more than six months behind the original schedule by next autumn, when the group should submit an engineering design report. "We expect to be there when the machine turns on." **Robert Pool**

ZOOLOGICAL RESEARCH

Zoo conservation plan

London

THE House of Commons Environment Committee has endorsed proposals by the Zoological Society of London to save the threatened London Zoo by transforming it into a centre for conservation education.

In this form, the zoo would display fewer animals and would emphasize the work of the society's research arm — and in particular, the Institute of Zoology.

The zoo has been losing money for years, and in 1988 the government gave the Zoological Society — which runs the zoo — a grant of £10 million that was supposed to put the zoo back on its feet. But optimistic predictions of an upsurge in visitor numbers were not fulfilled, and the society announced earlier this year that the zoo would close unless more money were found.

On 9 July, the society's trustees must decide whether to close the zoo immediately or to use the remaining £4.5 million from the

RESEARCH

Bush salutes NIH at Healy's swearing-in

Washington

BERNADINE Healy, the first woman to head the US National Institutes of Health, was officially sworn in this week at a ceremony with President and Mrs George Bush as the guests of honour. In a city in which symbols and reality are often one, the President's visit to the NIH campus in Bethesda, Maryland, on the outskirts of Washington, was meant to boost the morale of the nation's biomedical researchers, who feel mired in a financial crisis. The president called the NIH staff, from grants administrators to animal care staff, "unsung heroes" and said that "biomedical research is key to transforming the practice of medicine".

Healy, a Harvard cardiologist by training, also talked about the NIH's role in medicine. The purpose of all the fundamental work in molecular biology and other areas of basic research is "basically simple", she said. It is "saving lives". In emphasizing medical care as the ultimate goal of all research sponsored by the National Institute of Health, she reiterated a theme she struck during her confirmation hearings before the US Senate (see *Nature* 350, 178; 21 March 1991).

And she continued to present herself as a physician first and foremost with a story about a young woman in the late stages of metastatic breast cancer. Taking Healy's arm after a conversation about efforts in research, the woman said "Dr Healy, please hurry." Speaking to a crowd of several hundred guests at the swearing-in, Healy said, "I take that as my mandate."

Barbara J. Culliton

1988 grant to implement a 3-year "phased withdrawal" from the Regents Park site.

The first option implies destruction of many of the zoo's animals, which the committee condemns as "totally unacceptable".

The second, says zoo director David Jones, would imperil the future of Whipsnade, the society's other zoo site, outside London, where some of the Regents Park animals would be rehoused.

If the society is successful in winning private sponsorship for a conservation centre, the House of Commons committee concluded that the government should provide some money to help implement the scheme. The committee noted that if the society were to close the Regents Park site, the government would still end up paying for it — the Department of the Environment, which leases the grounds to the society, could be faced with a bill of several million pounds to redevelop the site.

Henry Gee

AIDS threatens Asia

Florence

New data reported last week by experts at the 7th International AIDS Conference in Florence indicated that developing countries are now facing an even worse AIDS emergency than previously believed.

The latest epidemiological studies performed by the World Health Organization (WHO) show that even though the incidence of AIDS is still low and is just beginning to be reported in many developing countries such as India, the disease is beginning to spread at an alarming speed. Although earlier data had shown Africa to be seriously endangered by the spread of HIV, the virus that causes AIDS, the new data reveal an alarming rate of increase in both Asia and Latin America as well.

By 2000, more than 90 per cent of HIV infections in the world will probably be found in the developing world, reports James Chin, a WHO epidemiologist in Geneva. At the same time, the number of new infections in industrialized countries will begin to level off or even to decline because of changes in behaviour.

The new data show that HIV has spread in India far beyond its initial entry points of the port cities of Madras and Bombay and is beginning to threaten rural populations, which are virtually defenceless against the virus because of a lack of information and condoms. One Indian official, Vulimiri Ramalingaswami of the All India Institute of Medical Sciences in New Delhi, said that India is "sitting on top of a volcano".

In Thailand, HIV has begun to race

through the population with a remarkable speed, says Chin. The prime minister of Thailand recently told WHO that the estimated number of people infected with HIV increased from 200,000 to 400,000 in just the six months between June and December 1990. The prevalence of HIV among male military recruits had jumped from two per cent to six per cent during the same period.

Accordingly, WHO raised its estimate for the present number of HIV infections in South Asia from 500,000 to 1,000,000, Chin said. While he stressed a basic uncertainty about the figures (which other epidemiologists said could be off by as much as 95 per cent), he also said that it is more likely that the WHO estimates will be revised upward rather than downward as more data become available.

At the same time, there is little money available in many developing countries to combat AIDS. In Uganda, for example, President Yoweri Kaguta Museveni said that per capita health spending amounts to just \$3.50 a year — not nearly enough to treat people with AIDS, let alone try to prevent infection.

By 2000, about 90 per cent of the transmissions of the virus are expected to occur through heterosexual intercourse, despite the fact that this is one of the least efficient methods of transmission. Therefore, William Haseltine of Harvard University said it is time that AIDS, originally thought of as primarily a homosexual disease, should be recognized as being a "lethal venereal disease" that can be transmitted from fully

healthy men to fully healthy women through mucous membranes.

■ Meanwhile, the dispute over who discovered the AIDS virus flared up anew, despite unprecedented calls from activists asking the principals to stop wasting their time on the matter.

Luc Montagnier of the Pasteur Institute, officially recognized as the co-discoverer of the virus and its role in causing AIDS, along with Robert Gallo of the US National Institutes of Health, told *Nature* that "France deserves a gesture from NIH and our American colleagues" to make up for time and money wasted by the dispute "because of the early fault of Dr Gallo. There should be some reparations" to make up for that, he said.

Montagnier and the director of Institut Pasteur, Maxime Schwartz, denied that Institut Pasteur is interested in re-opening the arguments over the royalties from an AIDS blood test developed in Gallo's laboratory — royalties which have been split 50–50 since an agreement in 1987 between NIH, Institut Pasteur and the two governments involved.

But Schwartz hinted that if Institut Pasteur is displeased with the result of a current NIH inquiry into the Gallo laboratory's handling of the case, then the Institut Pasteur might choose to do something about it. "This is not over yet," he said.

Meanwhile, Project Inform, an AIDS activist group based in San Francisco, urged an end to what it called a "harassment and misinformation campaign" against Gallo. It called on Gallo, Montagnier, *Chicago Tribune* journalist John Crewdson and others in any way involved in the dispute over the origin of the virus to "end this ludicrous debate".

Steven Dickman

Researchers bow to activist pressure

UNDER heavy pressure from US activists, the chairman of next year's international AIDS conference last week reversed his position and announced that the conference will not be held in Boston in 1992 unless the United States drops its restrictions on the entry of both travellers and immigrants who carry HIV, the virus that causes AIDS.

Max Essex of Harvard University said he changed his position after receiving "ultimatums" from activists who threatened to disrupt the conference if it were held. Activists were reported to have said that "blood would run in the streets" of Boston if the meeting were held.

If the meeting is not held in Boston, however, chances are that it will not be held at all. The International AIDS Society, which has been the official sponsor for each of the first seven international AIDS conferences, announced here that if the conference could not be held in Boston, the society would not sanction any other site. The next conference after that is scheduled to be held in Berlin in 1993.

For the meeting to be held in the United States, the US government would have to reverse its position again on the issue of travel and immigration restrictions — an occurrence that US government researchers like Anthony Fauci of the National Institutes of Health considered "almost

inconceivable". Essex said that unless all US travel and immigration restrictions against HIV-infected people are lifted by 3 August, "there will be no conference in Boston".

At this early stage, it is hard to predict the consequences of Essex's switching his position under pressure. But Meinrad Koch of Germany, executive secretary of the International AIDS Society and a co-organizer of the 1993 Berlin meeting, criticised Essex for "going overboard" in demanding that the immigration restrictions be reversed. "He was playing to a vocal minority," said Koch.

Koch sees two possible outcomes from Essex's decision, which he characterizes as a case of science bowing before irrationality. Either "scientists will become prisoners of the activists," he said, or there will be increasing hostility among scientists toward the activist groups. Either way, it will contribute to making the Berlin meeting a difficult task in a city which is the home of the largest gay community in Germany.

Steven Dickman



Demonstrating against discrimination.

SOVIET SCIENCE

Rent a theorist in Moscow

Washington

SOVIET physicists, fearful of the impending economic 'anarchy' in their country, have dug deep into the capitalist bag of tricks to find a way to keep science alive in the USSR: they propose to rent themselves out.

For a price of between one-quarter and one-fifth of what research costs in the West, Soviet theoreticians at the Kurchatov research institute claim they can do the same or better calculations in Moscow. So they have set up an independent group of between 100 and 200 physicists, secured some Soviet government funding, and spread the word to their Western contacts that they are open for business for anyone who needs help with tough calculations.

Known as Mucatex, the group is headed by Kurchatov theoretical physics director Leonid Ponamarev and is said to include some of the best theoretical physicists in the world. Their speciality is muon-catalysed fusion, the technique of exchanging a heavy muon for the electron in a deuterium atom, which allows two deuterium atoms to get close enough to occasionally fuse and release large amounts of energy.

Although muon-catalysed fusion fell somewhat out of favour in the United States recently after being tarred by the 'cold fusion' brush, it is still an active field of research in dozens of laboratories worldwide. Supporters say it could play a critical role in future fusion technology by enhancing the performance of breeder reactors by a factor of two or three.

Western contacts for the Soviet researchers say their proposition is an attractive one for both sides. Researchers in the West can get time-consuming calculations done by

some of the best theoreticians in the world at a cost far less than in the West. The Soviet government would get some hard cash and, more importantly, would be able to keep their top scientists from leaving the country. Organizers estimate that each deal would include about 20 per cent travel time for the Soviet researchers to visit the West to work with their sponsors.

"It took me 10 to 15 years to create a team that includes elementary-particle, quantum-physics, computer and accelerator physicists," says Ponamarev. "Now I am worrying that it will disappear in the coming chaos. To survive this group we need outside support." Ponamarev hopes that some of the time his researchers will spend in the West will be on high-performance computers. Computer resources are scarce in the Soviet Union, and Western nations are often reluctant to allow Soviet institutions direct access to their supercomputers, which are often used for classified research. "They'll probably do their thinking in Moscow and their number crunching in the West," says Birmingham University physicist John Davis, who has been in contact with the Soviet group.

Working under the umbrella of the World Laboratory (based in Lausanne, Switzerland), Mucatex has already established collaborations with the Paul Scherrer Institute in Zurich, the UK Rutherford Appleton Laboratory, the US Los Alamos Laboratory and institutions in Italy and Germany. The group is currently circulating a letter by Soviet Academy of Science Academician Eugenu Velikhov, seeking further collaborations. Western interest so far has been encouraging, says Ponamarev.

Christopher Anderson

HUMAN GENOME PROJECT

New office for HUGO in Soviet Union

Washington

THE Human Genome Organization (HUGO), the body set up to help coordinate the international effort to map and sequence the human genome, is opening an office in Moscow. The aim, says HUGO vice-president Charles Cantor from Lawrence Berkeley Laboratory, is to improve the lines of communication between genome researchers in the Soviet Union and those in the rest of the world.

In terms of manpower, the Soviet genome project is second in size only to that of the United States. But poor communication links with the West hamper coordination between the Soviet programme and the other national genome efforts.

The Moscow office will open on 1 July, and Cantor expects electronic communications to be established via a satellite link within several months. This should give Soviet genome researchers rapid access to

the computer databases in which gene mapping and sequencing data are being accumulated. The costs of the plan have not yet been estimated, but Academician Nikolay Laverov, Soviet deputy prime minister, told HUGO's executive committee in Moscow last week that the Soviet authorities will provide most of the money.

The office is also expected to become the focus for internal coordination of the Soviet genome project. The Soviet pledge of financial support was welcome — HUGO's international coordinating role has so far been limited by its small budget.

The new office will be HUGO's fourth, adding to offices in London and Maryland, and a Japanese office in Osaka, still in the process of being set up. HUGO staff describe the Moscow office as a "satellite" of their office in London, which will continue to be the focus of HUGO's activities in Europe.

Peter Aldhous

JAPAN

Human Frontiers gets extension

Japan

JAPAN'S unique Human Frontier Science Program (HFSP), which supports international research on the brain and molecular biology, seems set to continue until at least 1997. That is the most important decision to have arisen out of a HFSP board of trustees meeting earlier this month in Strasbourg, France.

The board of trustees, which consists of representatives from the seven Western summit nations (Japan, the United States, Canada, France, Italy, Germany and the United Kingdom) as well as the European Community and Switzerland, unanimously agreed at the closed meeting on 5–6 June to continue the programme for another five years after the initial three-year phase ends in March 1992. At the end of the next five-year phase, Frontiers will undergo major review with the possibility of extension for another five years, Japanese government officials say.

A confidential report from the meeting, which is being circulated among various ministries in Tokyo, also says that the board expects financial contributions from other member nations apart from Japan to increase "until an equal match is achieved". At present, Japan is providing about 80 per cent of the programme's annual budget of approximately \$30 million.

A budget proposal for fiscal 1991 presented at the meeting reveals that Germany has agreed to make a substantial "in-kind" contribution to the programme, paying more than \$1 million to cover the salaries of HFSP fellowship awardees. This will make Germany the third largest contributor after Japan and France.

The proposed budget also includes an in-kind contribution from the United States of about \$40,000 to support the salary of one HFSP fellow at the US National Institutes of Health. This confirms earlier reports that the institutes are prepared to make a tiny but politically significant contribution to the programme (*Nature* 350, 97; 1991). This leaves the United Kingdom as the only member state that is not making any financial contribution to Frontiers.

Japanese government officials say they are generally pleased with the outcome of the Strasbourg meeting and with the growing financial support for the programme from other member nations, but there is one matter that irritates them. France is putting a 24 per cent tax on interest earned from HFSP funds deposited in France, where the HFSP organization is based. This tax nets the French government nearly \$750,000, or almost half of the country's \$1.5-million annual contribution to Frontiers. Japan and other member nations have appealed to France to exempt the programme from tax, but so far to no avail.

David Swinbanks

Congress reviews DNA testing

Washington

BILLS to regulate the forensic use of DNA fingerprinting are expected to be introduced into both houses of Congress this week.

The proposed legislation aims to address widespread concern, expressed by defence lawyers and some academic biologists, about the quality of DNA evidence now coming before US courts.

The bills, drafted by Don Edwards (Democrat, California) in the House and Paul Simon (Democrat, Illinois) in the Senate, would make federal funds available to state and local crime laboratories conducting DNA testing. The director of the Federal Bureau of Investigation (FBI), advised by a permanent independent advisory panel, would draw up standards for DNA testing, and those standards would bind the central forensic laboratory of the FBI as well as any laboratories that take federal funding. The proficiency of each scientist producing DNA fingerprints in these laboratories would be tested at least twice a year.

At present, there are no legally binding standards to assure the quality of DNA fingerprinting evidence, which is now being produced by the FBI, 13 state and local laboratories and two main commercial DNA testing laboratories, Cellmark Diagnostics of Germantown, Maryland and Lifecodes, of Valhalla, New York.

The proposed legislation would not apply to the commercial laboratories, nor to any state and local laboratories choosing not to apply for federal funds. But once standards are published, courts are likely to demand that any laboratory submitting evidence obey the standards and subject itself to the same proficiency testing proposed for government-funded crime laboratories.

The need for published standards is generally accepted by forensic DNA testing practitioners. Kevin McElfresh, director of identity testing at Lifecodes, says the lack of standards can be used by defence lawyers as a weapon to challenge DNA evidence.

Although DNA evidence has been accepted by most courts in which it has been presented, DNA fingerprints have been ruled inadmissible in a handful of celebrated cases. But biologists who have testified as expert witnesses for the defence in cases involving DNA evidence argue that tighter regulation is needed for a more sinister reason — the possibility that sloppy laboratory work could result in the innocent being convicted of murder or rape, or the guilty going free. Robyn Nishimi, the author of a congressional Office of Technology Assessment report on forensic DNA testing published last year (see *Nature* 346, 499, 1990), agrees: "It's not just an admissibility in court issue, it's a quality of data issue." Problems can arise in DNA fingerprinting when initial digestion of DNA using restriction enzymes is incomplete, or if the electrophoresis condi-

tions used to separate the DNA fragments vary.

These errors can often be spotted, but simpler mistakes — such as the mislabelling of sample tubes, or cross-contamination from one sample to another — can slip through unnoticed. Simon Ford, a California-based biotechnology consultant who has acted as an expert defence witness in cases involving DNA testing, points to the results of one of the few published DNA fingerprinting proficiency tests. In 1987 and again in 1988, the California Association of Crime Laboratory Directors sent Cellmark 50 samples to fingerprint and to group those that matched. On both occasions, Cellmark came up with one false match.

Many population geneticists are critical of the methods used by forensic laboratories to calculate the probability of an observed match between DNA profiles occurring by chance. Dan Hartl, from Washington University, St Louis, believes that the probabilities quoted in court "can easily be off by a factor of a thousand". The problem is that forensic scientists usually ignore the substructuring within Caucasian, black and Hispanic populations. Within population subgroups, band sharing on DNA fingerprints may be much more common than for the population as a whole.

Edwards' draft bill calls for a permanent independent advisory board, including molecular geneticists, population geneticists and legal experts, to be set up by the National Institute of Standards and Technology to develop standards for forensic DNA testing and to advise the director of the FBI. The involvement of academics and lawyers is not to the liking of all forensic scientists. But Eric Lander, from the Whitehead Institute at the Massachusetts Institute of Technology, who has taken a high profile in the debate over forensic DNA testing (see *Nature* 339, 501, 1989), says that academic participation is essential. "This is an issue of the transfer of technology from academia to forensics," he says. The bill also contains clauses to ensure that DNA test data identifying an individual are obtainable only by criminal justice agencies and the defendant.

In drafting their bills, Edwards and Simon have jumped the gun on a National Academy of Sciences report on forensic DNA testing, due later this summer. But congressional staff have been consulting with the academy, so there should be no major conflicts between the bills' provisions and the academy's recommendations. A far less extensive bill on forensic DNA testing was introduced into the House earlier this year by Frank Horton (Republican, New York). But as Edwards and Simon are chairmen of House and Senate subcommittees that are responsible for the application of forensic science, the new bills should carry more weight. **Peter Aldhous**

Closing a loophole in discrimination rules

Washington

REGULATIONS due to be published next month could allow US employers to discriminate against job applicants on the basis of their genetic make-up. To avoid this possibility, Nancy Wexler, who chairs the US genome project committee dealing with the ethical, social and legal implications of genome research, will this week ask the National Institutes of Health (NIH) to press for changes to the proposed rules.

The regulations in question were drafted earlier this year by the Equal Employment Opportunity Commission (EEOC). Their purpose is to implement the 1990 Americans with Disabilities Act — legislation supposed to outlaw discrimination against job applicants on grounds of disability.

The EEOC regulations state that employers should not decide whether to hire a person on the basis of medical tests that have no bearing on that person's ability to do the job. But the regulations, as they were drafted, would allow employers to conduct tests that are not job-related, once a conditional offer of employment has been made. Because employers are not bound to reveal which medical tests they have done or why an offer of employment has been withdrawn, the regulations could do nothing to stop employers determined to discriminate.

This is a major concern in the case of genetic screening, because it may be possible in future to identify a whole suite of genes that either cause disease, or confer a high susceptibility to particular health problems. The worry is that, to minimize their health care provision costs, some employers will screen for these genes in order to exclude job applicants who are more likely to become ill. "The regulations open up some loopholes that were closed by the Act," says Eric Juengst, from the National Center for Human Genome Research.

Wexler, who is at the Columbia College of Physicians and Surgeons, wants employers restricted to conducting medical tests that are directly job-related, and hopes to convince the NIH to use its weight to secure changes to the EEOC regulations at this week's meeting of the US genome project's advisory committee.

The final regulations will be published next month, but will not come into force for another year. Mark Rothstein, director of the Health Law and Policy Institute at the University of Houston, says that when the Act was passed, there was little debate of the future impact of genetic testing. "People are only now beginning to think about it," he says. This will be the first time that Wexler's committee has asked the NIH to intervene in the affairs of another Federal agency. **Peter Aldhous**

Clearing-up the Gulf

SIR — I would like to add to the views expressed on the systematic dismantling of Kuwait University facilities. I had been an academic staff member of the faculty of medicine for more than six years; for a part of the time after the invasion, I was acting chairman of the physiology department.

During last summer, in the absence of the dean, the acting dean of the faculty of medicine was an Iraqi. After the invasion, he organized the closure of the faculty, taking all measures possible to safeguard its facilities. He remained loyal to the faculty and to the Kuwaiti government at great personal risk; finally he decided to escape from the Iraqi authorities in September.

I can sympathize with the points expressed by Abdalla O. Elkhawad (*Nature* **351**, 9; 1991). To have played a role in the growth of that institute and then to see it destroyed is heart-rending. Unfortunately, evidence suggests that Iraqis found collaborators prepared to change sides. I know from my own experience that some expatriates who joined at the inception of the faculty did not strive for academic achievement alone. They had their own vested interests at heart and gained not only financial reward but also power and high office, and consequently gave undue advantages to their kinsmen.

One therefore wonders how these expatriates aligned themselves after the Iraqi invasion. The initial assessment was that the new rulers were to be Iraqis. What role would these expatriates have then played since their vested interests were at stake, and their governments had already aligned with Iraq? Did they play any part in the systematic dismantling of Kuwait University?

It is inconceivable that without inside informers the Iraqi authorities could within a short time have known of the location and content of sensitive equipment and undertaken its dismantling and relocation. Could it be that the role of builders became the role of destroyers — as evidenced by reports of collaboration from Kuwait?

B. AHMED

MRC Anatomical
Neuropharmacology Unit,
University of Oxford,
Oxford, OX1 3QT, UK

SIR — I share Samir Radwan's concern¹ that the useful contribution of bioremediation technology to the treatment of the Alaskan oil spill may cloud our view of how best to deal with the Gulf spill.

It is not surprising that his research group found rhodococci to be the predominant oil-degrading bacteria in the Kuwaiti environment². These organisms belong to the 'nocardioform' group of actinomycete bacteria, renowned for their ability to utilize hydrocarbons as growth substrates. They are also better equipped to survive desiccation, and thus persist in arid environments, than

the more familiar groups of unicellular, oil-degrading bacteria. Nocardioform actinomycetes have a lipid-rich cell envelope, and difficulties in controlling their growth have been encountered before.

In sewage treatment plants, separation of biomass from purified effluent can be seriously impeded by the formation of a network of actinomycete hyphae which entrap and feed upon oil and grease. The phenomenon is referred to as 'sludge scumming' or 'nocardial foaming'³ and large amounts of microbial biomass in this form, rather than as sedimentable flocs, would add an extra dimension to the problem of processing water supplies drawn from the Gulf.

A. J. MCCARTHY

University of Liverpool,
Department of Genetics and
Microbiology,
PO Box 147, Liverpool L69 3BX, UK

1. Radwan, S. S. *Nature* **350**, 456 (1991).
2. Sorkhoh, N., Ghannoum, M. A., Ibrahim, A. S., Stretton, R. J. & Radwan, S. S. *J. appl. Bact.* **69**, 856–863 (1990).
3. Lemmer, H. in *Biological, Biochemical and Biomedical Aspects of Actinomycetes* (ed. Szabo, G. et al.) (Akademiai Kiado, Budapest, 1986).

SIR — I am disturbed by the speculation of S. C. Coburn (*Nature* **351**, 9; 1991) that a bombed milk production facility could have been the source of media for the cultivation of military microorganisms. Where is the evidence that anything like this was going on? We do not require any further *post hoc* explanations of what we did in the Gulf just to salve our consciences. The systematic destruction of the infrastructure of Iraqi society by strategic bombing was an outrage against humanity at least as abhorrent as the invasion and annexation of Kuwait by Saddam Hussein's forces.

PETER R. WILLS

Physics Department,
University of Auckland,
Auckland, New Zealand

■ Climate effects of the Gulf oil spill, see page 704.

Keeping cool

SIR — I should like to clarify the remark attributed to me in your news item "Keeping Cool in Space" (*Nature* **351**, 338; 1991).

Although the lifetimes of the Infrared Astronomy Satellite and the Infrared Space Observatory missions can be conveniently expressed in months (10 and 18+ respectively) rather than years, this should not be taken to imply that cryogenically cooled missions cannot be engineered to operate for longer periods. The NASA Space Infrared Telescope Facility, which will carry a cryogenically cooled 1-metre class telescope, is expected to operate for five to six years before its helium coolant is exhausted.

JOHN DAVIES

Royal Observatory, Edinburgh, UK

Reduced circumstances

SIR — Douglas B. Kell (*Nature* **350**, 268; 1991) bemoans the lack of knowledge of the term 'reductionism' among his first-year biology students. Presumably they receive no formal tuition on this subject in his department's teaching programme, but does he expect schools to be teaching reductionism as a philosophical idea?

At no time since I began A-level biology teaching in 1968 have I been required to teach reductionism to fulfil formal syllabus requirements, to fall in with the suggestion of more senior colleagues or even to comply with my own judgement of what students should be taught.

In my department we use two A-level texts^{1,2}. Neither mentions reductionism in the index or in the introductory chapter, I cannot recall seeing the word in texts aimed at sixth-form level, and those more advanced texts that we do use for reference^{3–5} contain no mention of reductionism. In the circumstances, does not responsibility lie with universities and polytechnics?

ROLAND DIXON

Biology Department,
Wellington School,
Wellington,
Somerset TA21 8NT, UK

1. Green N. P. O. et al. *Biological Science* Vols 1 & 2 (Cambridge University Press, 1990).
2. Roberts, M. B. V. *Biology – A Functional Approach* 3rd edn (Nelson, Walton-on-Thames, 1982).
3. Watson, J. D. et al. *Molecular Biology of the Gene* 4th edn (Benjamin/Cummings, Menlo Park, 1987).
4. Darnell, J. et al. *Molecular Cell Biology* (Scientific American Books, New York, 1986).
5. Alberts, B. et al. *Molecular Biology of the Cell* 2nd edn (Garland, New York, 1989).

Animal welfare

SIR — I appreciated Christopher Anderson's coverage of the US court ruling that paves the way for a legal challenge to the US Department of Agriculture (USDA)'s exclusion of mice, rats and birds from the protection of the Animal Welfare Act (AWA) (*Nature* **350**, 642; 1991). However, two clarifications are needed.

The Animal Legal Defense Fund was identified as the plaintiff in the suit against the USDA but other plaintiffs include the Humane Society of the United States (HSUS) and two individuals. Also, the report stated that the USDA has the authority under the AWA "to determine what is and is not to be defined as a research animal".

Whether or not the USDA has this authority is the core issue in the suit. The plaintiffs believe the USDA is exceeding its authority by excluding mice, rats and birds from AWA regulations.

MARTIN L. STEPHENS

The Humane Society
of the United States,
2100 L Street, NW,
Washington, DC 20037, USA

Evolution on the mind

SIR — Although I — probably like many of your readers — do not quite see why Waller's five-wish 'evoholics' should be so different from Dawkins' selfish geniuses in practice (since the selfish gene model in no way precludes the evolution of genes for each of the five 'wishes' Waller lists), it nevertheless seems worthwhile to point out that psychoanalytical insights might be added to the list of the "various research findings" which Waller claims have "shown that all the 'gifts' which I have hypothetically given to my heirs actually exist" — at least among human beings¹.

For example, the unconscious sexual and aggressive wishes uncovered by psychoanalytical investigations in general, and freudian dream-interpretation in particular, might easily follow Waller's first wish (continual assessment of performance regarding genetic replication), and his second (a sense of well-being as a motivator, especially in stimulating "libido" — Waller's own term), seems to be Freud's "pleasure principle"^{2,3}. Psychoanalytical insights into mania as a defence against depression would certainly illustrate his third wish (reactive responses to depression), and his fourth (self-elimination in the event of failure) is perhaps the real biological foundation for Freud's hypothesized — and, until now, seemingly biologically absurd — "death instinct"⁴.

As far as Waller's final wish (a genetic basis for differing character types) is concerned, freudian findings regarding the importance of childhood identifications may provide illustrations in profuse detail. To take just one example, the widely reported finding from both conventional, academic psychology⁵ and psychoanalysis that fathers appear to be critical in determining the sex role of sons may in reality reflect a specific example of this fifth wish: one in which sons estimate the likelihood of their having inherited genes for a successful male sex role by means of identification (that is, phenotypic matching by psychological self-comparison) with an adequate father-figure, or otherwise as the case may be⁶.

Although Waller may well be as reluctant to recognize such freudian proclivities in his hypothetical offspring as parents frequently are to admit comparable evidence of genetic self-interest in themselves and their children, Freud for his part would probably have claimed honorary membership in the guild of selfish-gene psycho-darwinists on the strength of his assertion that "the individual organism, which regards itself as the main thing and its sexuality as a means, like any other, for its own satisfaction, is from the point of view of biology only an episode in a succession of generations, a short-lived appendage to a germ-plasm endowed with immortality — like the temporary holder of

an entail which will outlast him"⁷.

Indeed, Freud's real offence may have been to discover and describe the actual manifestations of this in human beings, rather than merely to speculate in general terms as Waller and others have done.

CHRISTOPHER BADCOCK

Department of Sociology,
London School of Economics
London WC2A 2AE, UK

1. Waller, M. J. C. *Nature* **351**, 264 (1991).
2. Barash, D. *Sociobiology: The Whisperings Within* (Souvenir Press, London, 1979).
3. Badcock, C. R. *Evolution and Individual Behavior* (Blackwell, Oxford 1991).
4. Badcock, C. R. *Essential Freud: An Introduction to Classical Psychoanalysis* 2nd edn (Blackwell, Oxford, in the press).
5. Lamb, M. *The Role of the Father in Child Development* (Wiley, New York, 1981).
6. Badcock, C. R. *Oedipus in Evolution: A New Theory of Sex* (Blackwell, Oxford, 1990).
7. Freud, S. *Introductory Lectures on Psychoanalysis*, **XXV**, 413–414. (Hogarth Press and the Institute of Psychoanalysis, London, 1963).

SIR — I found M. J. C. Waller's psycho-darwinist hypothesis of evolution very sobering, particularly in the light of my scientific interests as a pharmacologist working on the central nervous system. It appears to me that all we need do is wait for mother nature to eradicate CNS disorders of mood and eventually put the likes of me out of business. In a more serious approach to this theory, I feel a little concerned for the future of mankind. Many artists seem to have suffered from some kind of affective disorder, which would eventually push them out of the gene pool, and leave the remaining happy humans without much in the way of emotional and spiritual expression. It also occurs to me that, in most cases, depressive illness, like many other CNS disorders, appears to strike after most people have had their heirs (median age for depression is 40, and for bipolar disorders about 30). Thus, the psycho-darwinist theory, if it were true, would have little influence on the evolution of mankind.

M. S. FAZELI

Royal Free Hospital School of
Medicine,
London NW3 2PF, UK

SIR — M. J. C. Waller's thought-experiment (he and I as the only males on an otherwise all-female planet) has undeniable appeal. But alas, like many such fantasies, it has no connection with reality.

He imagines that we are each offered the opportunity to force our remote descendants to behave in certain ways so as to improve the 'evolvability' of our posterity. He would eagerly seize the opportunity and would, for example, require his descendants to sacrifice themselves, if they sensed that they were unsuccessful, for the good of their 'kind'. I, on the other hand, should honourably decline the offer, "confident that the selfish gene model cannot be improved upon".

Of course the selfish gene model can be improved upon. It is easy to see that an all-powerful dictator, with the interests of some

larger community at heart, could improve upon it. You name an interested group — the Coast Redwood forests, the world's whale population, albinos — and I'll specify an 'improvement' to the selfish gene theory that would benefit that group. The problem is not that one can't think of improvements from some named point of view. The whole problem is that those improvements will not be evolutionarily stable. They'll disappear under natural selection because the only point of view natural selection cares about is that of the replicating entities — the selfish genes.

Waller correctly discerns that his thought-experiment has something to do with group selection. Unfortunately it solves none of the notorious problems of group selection. If the set of all Waller's descendants could somehow be delimited as a watertight group, he might (just) have some hope of making a form of group selection work. But of course his descendants on the planet, far from being watertight, would pretty soon also be my descendants. Neither Waller nor I would be entitled to label our remote descendants as our 'kind' rather than the other patriarch's. His genes and mine would shuffle and reshuffle down the generations. In such a world of reshuffling replicators, the only interests that natural selection can favour are the interests of the replicators themselves.

There is a realistic sense in which lineages may be selected for improved 'evolvability', but it is very different from Waller's sense. I have spelled it out in my paper 'The evolution of evolvability'¹.

RICHARD DAWKINS

Department of Zoology,
South Parks Road,
Oxford OX1 3PS, UK

1. Dawkins, R. in *Artificial Life* (ed. Langton, C.) 201–220 (Addison Wesley, Santa Fe, 1989).

Snow's way with cholera

SIR — It was indeed John Snow who, in August 1854, investigated the cholera epidemic in the area around Broad Street¹. He later wrote: "The result of the inquiry then was, that there had been no particular outbreak or increase of cholera, in this part of London, except among the persons who were in the habit of drinking the water of the [Broad Street] pump-well . . . I had an interview with the Guardians of St James's parish, on the evening of Thursday, 7th September . . . In consequence of what I said, the handle of the pump was removed the following day."²

BRIAN MOONEY

University of Georgia,
Savannah River Ecology Laboratory,
Aiken, South Carolina 29802, USA

1. Bulloch, W. *The History of Bacteriology* 167–168 (Oxford University Press, London, 1938).
2. Snow, J. *On the Mode of Communication of Cholera*. Reprinted in *Source Book of Medical History* (ed. Clendening, L.) 468–473 (Dover, New York, 1960).

Analysis of a whistle-blowing

Walter W. Stewart and Ned Feder

The following article, published for completeness, was completed on 30 September 1987, was submitted to *Nature* soon afterwards, but never published. Those consulted about the manuscript held that there might be data, other than that in the seventeen pages of laboratory notebooks analysed by the authors, that would undermine their conclusions. Such supplementary data have indeed been submitted, but have been judged by the draft report of the Office of Scientific Integrity to have been fabricated.

IN the late spring of 1986, we were informed by an uninvolved party that the scientific validity of parts of a recent and important paper in molecular biology had been challenged by Dr Margot O'Toole, a postdoctoral fellow in the laboratory of one of the authors of the paper. We were told that the question was currently under investigation at two universities. Several months later we learned that neither university had taken public action.

Acting entirely on our own initiative, we contacted Dr O'Toole and asked what evidence she had to support her challenge. . . . She said she based her objections in part on discrepancies between the published paper and certain original data from the laboratory of one of the authors — data to which she had been given access during the course of her work. Dr O'Toole further informed us that she had raised these issues repeatedly with certain of the authors and with representatives of the two universities. She had been told that the presentation of data in the published paper conformed to accepted standards for publication of scientific results. . . .

We have a long-standing interest in the accuracy of the scientific literature and the efficacy of the processes that correct scientific error. It seemed likely that this paper¹ would be widely used and cited, as indeed it has been. . . . we decided to analyse the laboratory records ourselves in our capacity as scientists. We asked for and obtained from Dr O'Toole a copy of the laboratory records on which she based a part of her objections. We received from her 17 pages of laboratory records, which, together with the authors' replies to us and the report from one of the universities, form the basis of the analysis presented here.

The paper we analysed examines the repertoire of immunoglobulins produced by certain transgenic mice. The authors used a strain of C57BL/6 mice derived from an early embryo injected with a rearranged μ heavy chain gene from another strain². This gene, containing the joined V_H , J_H4 and C segments, had been incorporated into the germ line DNA of the transgenic mice. The transgene is expressed in the nonimmunized mice and produces a functional antibody². The idotype of the antibody has been designated 17.2.25. Using a polyclonal guinea pig anti-idiotype antiserum, the authors claimed that nonimmunized mice carrying the transgene had elevated levels of antibodies with

the 17.2.25 idiotypic.

One of the paper's central claims is that the elevated levels of 17.2.25 idiotypic seen in the nonimmunized transgenic mice are due in large part to endogenous antibody rather than to the product of the transgene. In particular the authors stated that the "various tests allow estimates that no more than 1–10% of the 17.2.25 idiotypic-positive Ig actually arises from the transgene".

According to the report¹, fewer than 1% of the hybridomas from nonimmunized normal C57BL/6 mice produced 17.2.25 idiotypic-positive antibodies. In contrast, a large fraction (172/340; 51%) of hybridomas from nonimmunized transgenic mice were reportedly idiotypic positive. Of the idiotypic-positive clones, a clear majority (76%) were reported to be producing antibody of endogenous origin.

In the authors' Table 3,34 transgenic hybridomas from a different fusion were characterized in greater detail. Only five of these were reported to express transgene. Most of the rest were said not to be expressing transgene on the basis of an S1 nuclease protection assay, and hybridization of northern blots with a 233-bp restriction fragment internal to the V_H gene.

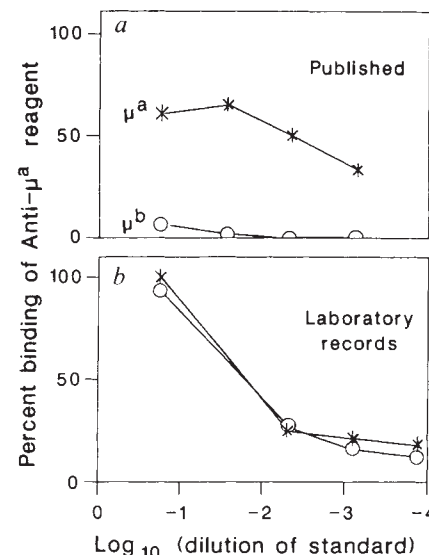


FIG.1 Discrepancy between published and laboratory data on Bet-1 reactions. The top panel shows the data published, the bottom panel the laboratory records. The μ^a standard (*) was the hybridoma 20 $\times$ 1 $\times$ 21, BALB/c ($\mu^a\lambda^21_2$). The μ^b standard (o) was the hybridoma P9.37.9, C57BL/6 ($\mu^b\lambda^21_2$).

The results appear to demonstrate unequivocally that the presence of the rearranged Ig gene in the germ line of the mouse in some way caused a marked expansion in the number of endogenous 17.2.25 idiotypic-positive clones. It is precisely this conclusion — a central finding of the paper — that is called into question by our analysis of the laboratory records.

As shown below, in some cases the laboratory data available to us clearly had been used in the preparation of the paper¹. In other cases, certain data were not published, and these data appear to have a direct bearing on the validity of the paper's conclusions. Despite repeated requests, the authors have not answered our specific questions about the data, nor have they made available additional laboratory data on which the paper was based. . . . In repeated responses, the authors have indicated to us that the presentation of data in their paper falls within the norms accepted by the scientific community and that the accuracy of their paper is a matter of interpretation. . . . In one case we were able to examine the report prepared by one [university] representative. The reservations presented below are those that remain after all parties concerned have had the opportunity, over a period of more than eight months, to respond to any or all of the issues we raised.

The 17 pages of laboratory records are essential for the analysis given below. The relevant authors have acknowledged that the records we examined had been brought to them by Dr O'Toole and that these records originated in the laboratory of one of the authors. Because of the importance of these records to our analysis, we describe them briefly in the Appendix [not reproduced here].

Specificity of Bet-1

The paper in question makes extensive use of the monoclonal reagent Bet-1 to test for exogenous μ^a produced by expression of the transgene. The paper repeatedly states that the Bet-1 reagent reacts only with exogenous μ^a and not with endogenous μ^b .

The data in the laboratory records show that the published statements are wrong. Data from these records, plotted in the bottom panel of our Fig. 1, show that Bet-1 reacted with both μ^a and μ^b proteins. . . . For comparison, the data published in the authors' Fig. 1 are reported in the top panel of our

figure.

We are well aware that standard curves such as those shown in the authors' Fig. 1 sometimes unexpectedly give anomalous results for unknown reasons. Such anomalous results on occasion have to be discarded as not representing the true reactivity of the reagent. We carefully considered whether this could be so for the data found in the laboratory records. Three arguments show that this is not the case. (1) The data actually published in the authors' Table 2 show several examples of this lack of specificity. Publication of these data by the authors shows they did not regard the observed lack of specificity as anomalous. . . (2) On the same day that the tests shown in our Fig. 1b were obtained, the reactivity with AF-6 and Bet-1 was measured for the 34 hybridomas reported in the authors' Table 3. This shows that the lack of specificity of the Bet-1 reagent was observed contemporaneously with measurements made on the hybridomas in Table 3. (3) Finally, the report from one of the universities states that the published description of the Bet-1 specificity is in error.

As shown in the next section, the laboratory records available to us contain the actual experimental data for the Weaver *et al.* Table 2. These data, which were actually published, confirm that in practice the Bet-1 reagent was not specific for μ^a protein. Measurements on a particular hybridoma, number 444 from transgenic spleen, illustrate this point. This hybridoma was reported in the authors' Table 2 as expressing endogenous μ^b . But the laboratory records show that this hybridoma gave more counts with Bet-1 (1,585 c.p.m., well above the cutoff value scored as positive by the authors, rising some days later to 7,775 c.p.m.) than with the authors' anti- μ^b reagent AF-6 (1,425 c.p.m.). These observations confirm that in actual laboratory experience, Bet-1 reacted not only with the μ^a allotype, but with μ^b allotype as well. However, in the authors' Table 2, this hybridoma was included among those producing endogenous μ^b . . .

The authors' publication of data from experiments where the Bet-1 reagent reacted with the μ^b protein shows that this reactivity was not regarded by them as an anomaly. It should be noted that under assay conditions different from those used by the authors, several groups^{3,4} have reported that Bet-1 reacts with the μ^a allotype to greater extent than with the μ^b allotype. The point we wish to make is simply that under the conditions used by the authors and according to their own data, the reagent did not in fact have the specificity that they explicitly stated it had.

The data published as their Table 2 showed a pattern of Bet-1 reactivity entirely different from the reactivity depicted in their Fig. 1. The claimed μ^a specificity of the Bet-1 reagent is crucial to the assertion, made repeatedly in the paper, that transgene expression did not occur in μ^b -producing hybridomas.

Thus, as employed by the authors, the Bet-1 reagent was not a specific anti- μ^a reagent. This result flatly contradicts the data they published as well as their published assertions concerning the expression of endogenous 17.2.25 idiotype. . .

Spleen hybridomas

The authors' Table 2 reports data showing that idiotype-positive hybridomas from the spleens of nonimmunized transgenic mice were common (28%). Data for the very same experiment are given in the laboratory records. That the published paper and laboratory records refer to the same experiment is shown by the exact agreement in seven numbers. . . . The actual data contained in the laboratory records demonstrate that the procedures chosen by the authors to analyse and present their data for publication were not appropriate.

It is evident from the distribution of radioactive counts that hybridomas in the authors' Table 2 were scored as idiotype positive if and only if the binding of the radiolabelled anti- κ plus anti- λ antibodies in the idiotype assay gave more than 1,000 c.p.m. But no counts or other quantitative data were published; in Table 2 the authors reported only that hybridomas either were or were not idiotype positive.

Their method of analysing the data are invalid for the following reason: the data provide no evidence for the existence of two populations based on idiotype expression. The use of a cutoff figure of 1,000 c.p.m. is arbitrary and inappropriate. A frequency histogram of the data shows that this cutoff interrupts a continuous distribution near the middle of a peak and is therefore invalid (see our Fig. 2).

Another problem with the authors' procedure became apparent when four clones from transgenic spleen were retested in the idiotype assay a few days later. The laboratory records show that the counts for all four clones increased; the average went from

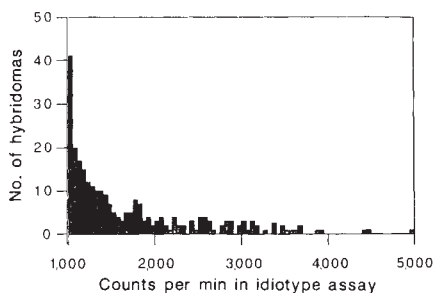


FIG.2 Radioactivity counts (per minute) in idiotype assay. The data shown are from available laboratory records; figures for normal and transgenic spleen and lymph nodes are pooled. Data from a total of 267 hybridomas are shown; values from a further 43 hybridomas were greater than 5,000 cpm and are not shown. Nor are the data for the remaining 382 hybridomas classified by the authors as negative, which were not recorded in the available laboratory notebooks.

6,912 c.p.m. to 28,698 c.p.m. This further emphasizes the arbitrary nature of the cutoff of 1,000 c.p.m. used by the authors. The observed increase suggests that if the test had been performed even a few hours later, more hybridomas would have been scored as idiotype positive. Moreover, the extreme time-dependency of the results of the idiotype assay on the length of time in culture suggests that the use of noncontemporaneous controls would not be appropriate. . . .

Lymph-node hybridomas

There are a number of peculiarities and inconsistencies in the original data for transgenic lymph nodes that show that the test results were scrambled or otherwise invalid. Despite these peculiarities, a summary of the data was published by the authors in their Table 2 with no mention of any possible defects.

The peculiarities in the data in the laboratory records were as follows. (1) Ten of the 129 idiotype-positive hybridomas were AF-6 positive; eight of the 10 were immediately adjacent to each other in the laboratory records, a striking, obvious, and remarkable coincidence. This obvious and unlikely clustering of the data clearly indicates a technical error. (2) For 18 of the 33 Bet-1-positive hybridomas, the Bet-1 counts were higher than the counts in the idiotype assay, a finding that is clearly anomalous. . . This again indicates a technical problem. (3) Seven of the 10 AF-6 positive hybridomas were Bet-1 negative, an anomalous finding that flatly contradicts the behaviours of the Bet-1 reagent observed and accepted by the authors and documented elsewhere in the laboratory records. (4) Two of the NP-positive hybridomas were producing both κ and λ , though this is not indicated in Table 2 for either of these hybridomas.

These four points indicate that the data were almost certainly scrambled. The laboratory records from subsequent experiments done several days later. . . emphatically confirmed that the original data were in error. Fourteen hybridomas were retested about 1 week later. Although nine of these had originally been Bet-1 negative, eight of the nine now tested Bet-1 positive. This showed that the original results were erroneous and further showed that at least eight additional hybridomas were actually Bet-1 positive. But only the initial results were reported in the authors' Table 2, without the inclusion of the eight additional Bet-1-positive hybridomas. The paper does not report that the typing of individual hybridomas with respect to expression of the transgene was time-dependent. Thus, in several ways, Table 2 presents the data incorrectly.

Just as in the data for transgenic spleen, counts in the idiotype assay for the hybridomas from transgenic lymph nodes increased substantially after several days in culture (average of 10,250 c.p.m. compared with 17,601 c.p.m.). The observed increase again shows that the authors' cutoff of 1,000 c.p.m.

was arbitrary and misleading. The authors' failure to report quantitative data, together with the methods they used to convert the numerical data into a simple assignment to just two categories, presents a picture favourable to the authors' conclusions, but one that was contradicted by the underlying data.

Hybridoma frequencies

An experimental finding essential to the authors' conclusions is that idiotype-positive hybridomas derived from nonimmunized normal mice are rare. Data to support this thesis are given in the authors' Table 2, which provide the expected result, namely, that such hybridomas are very rare (<1%).

The expected conclusion, however, is flatly contradicted by the authors' laboratory records for normal mouse control, which show that 138 of 352 hybridomas (39%) were observed by them to be idiotype positive in their assay system . . .

Thus the authors' assertion, central to the main conclusions of the paper, that hybridomas from normal mice are rarely idiotype positive, is contradicted by the laboratory records available to us. Despite repeated requests, the authors have not explained why they published a part of an experiment whose controls they discarded, nor have they explained why they discarded the test results of the experiments whose controls they apparently published.

In the appendix [not reproduced here] we consider the possibility that the data labelled "normal" in the laboratory records were actually obtained from hybridomas from transgenic mice. As discussed there, we feel that in any case these data, which represent the authors' true contemporaneous controls, are directly relevant to the truth of their thesis and . . . should therefore have been reported. Further, because the authors' method of classification depends strongly on the length of time the cells were cultured, the use of noncontemporaneous controls is not valid.

Antibody and expression

The authors' Table 2 presents an important finding. The majority (130 of 172) of the idiotype-positive hybridomas are shown as not expressing the exogenous μ^a transgene. This indicates a high frequency of hybridomas expressing endogenous idiotype-positive antibody. Further, only 11 were shown to be expressing endogenous μ^b . The authors assert: "Of the 172 idiotype-positive hybridomas, only 53 were IgM secretors". But the laboratory records contradict this claim.

In the authors' Table 2, hybridomas were reported in (+/-)-fashion as expressing μ^a only if a plate coated with the anti-idiotype reagent, followed by hybridomas supernatant, bound more than 1,000 c.p.m. when the radioactive Bet-1 reagent was added. Quantitative data were not reported; the hybridomas were reported simply as expressing μ^a or not.

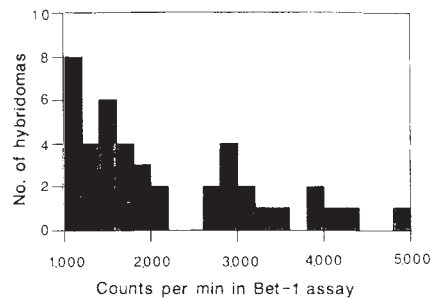


FIG. 3 Distribution of hybridomas in radioactivity (c.p.m.) for Bet-1 assay. Data from laboratory records of the Bet-1 assay used to score hybridomas as μ^a or μ^b . Data from 42 hybridomas are shown. Values for a further five hybridomas were greater than 5,000 c.p.m. and are not shown. The remaining 645 hybridomas classified as negative for μ^a were not recorded in the available notebooks.

The actual data contained in the laboratory records raise serious questions about the appropriateness of the reporting procedures . . . First, the quantitative data from the Bet-1 assay, presented in our Fig. 3, show that the criterion of 1,000 c.p.m. . . interrupts a continuous distribution near the middle of a peak and is therefore not valid. Second, . . . it is apparent from inspection of the graphed idiotype data (our Fig. 2) that there is nothing to support the use of a cutoff figure of 1,000 c.p.m. in determining whether or not a hybridoma is idiotype positive; indeed these data show that this cutoff was invalid as a means of discriminating between two populations.

Third, it is clear from the data graphed in our Fig. 4 that the Bet-1 counts are highly correlated with the counts in the idiotype assay ($r = 0.69$); with four outlying points removed, the correlation is excellent ($r = 0.94$). This high correlation suggests that for the great majority of clones, the idiotype reagent and the Bet-1 reagent are reacting with the very same substance, namely, the product of the transgene. This would explain the pronounced tendency for strongly idiotype-positive hybridomas to be μ^a -positive as well. . . . Thus the authors' assertion, made repeatedly, that most idiotype-positive hybridomas were not producing transgene protein is certainly not supported by the data. In fact, the data indicate the opposite.

Furthermore, according to the laboratory records, the Bet-1 counts for any given hybridoma are, with a single exception, less than the counts in the idiotype assay; the slope of the least-squares line is 0.29 (see legend to Fig. 2). The likely explanation is that the sensitivity of the former assay (Bet-1 binding) was reproducibly about 0.3 times that of the latter assay. In spite of this, exactly the same cut-off value, 1,000 c.p.m., was used to separate positive from negative in the two assays, despite the three-fold difference in their sensitivity.

The direct consequence of this invalid procedure was that a large number of transgenic hybridomas (our Table 2) were erroneously classified as μ^a negative and idiotype

positive (that is, producers of endogenous idiotype-positive antibody). We submit that there is no rational basis for this classification procedure.

A few days after the initial tests, a small number of clones were retested. For the 18 clones retested, the average of the Bet-1 counts rose from 2,137 c.p.m. to 8,071 c.p.m. This dramatic increase again shows the authors' classification system is vulnerable to differences in the length of time of culture; it shows that more clones would have been scored as μ^a -positive if a longer culture time had been used.

The authors' failure to report quantitative data, together with their use, for classification, of an arbitrary and time-dependent criterion, which they did not describe, is misleading. The laboratory records contradict the authors' assertion that endogenous antibody production accounts for the high frequency of idiotype-positive hybridomas. The striking discrepancy between the original laboratory data and the published results naturally raises the question: was the experiment that was published different from the experiment in the laboratory records? For the experimental data (but not the control data) shown in Table 2, there is compelling evidence that the experiments were the same . . . there is no reasonable doubt that both sets of data are from the very same experiment.

Other data

The laboratory records contain data on the isotypes, idiotype reactions and reactivity with the Bet-1 and AF-6 reagents for every one of the 34 hybridomas in the authors' Table 3. Although not stated in the paper, the data show that these hybridomas were produced at a much earlier time than the hybridomas shown in Table 2. The data show that at the time they were tested, 22 of the 34 hybridomas were Bet-1 positive (>800 c.p.m.). Since 21 of these 22 clones did not react with the AF-6 reagent, it is clear that they were expressing transgene. This is in flat contradiction to the authors' heading, "Most Transgenic Hybridomas Do Not Express the μ Transgene", which states one of the central conclusions of the paper.

Particularly disturbing is the contradiction between the original data and the published results for seven of the hybridomas in Table 3, namely, the seven hybridomas shown in the table to be producing μ^b and not μ^a . The laboratory records flatly contradict the results in Table 3. According to the laboratory records, five of these seven hybridomas reacted with the Bet-1 reagent and not with the AF-6 reagent. Thus for these five hybridomas, only the production of exogenous μ^a was detected.

There is further confirmation that many of the hybridomas were expressing the transgene. The laboratory records contain 5 instances where two heavy chain isotypes were recorded; in all five of these cases, one of the heavy chains was a μ . Without excep-

tion the authors reported just one isotype in their Table 3 (and in four cases they did not report the μ isotype) . . .

The paper in question states that most hybridomas did not express the transgene. In contrast, the laboratory records contain test results showing that most hybridomas were expressing transgene. These results were not included in the published paper.

We note that the authors, commenting on the 172 idiotype-positive hybridomas in Table 2, stated that there were 119 that did not secrete IgM. They further stated that these 119 clones "produced other Ig heavy chain isotypes, the majority being γ_{2B} (data not shown)." The laboratory records available to us contradict the assertion about the non-secretion of IgM, and fail to support the assertion about the γ_{2B} isotypes. Our repeated requests for data to support the assertion . . . or for an explanation have not been granted.

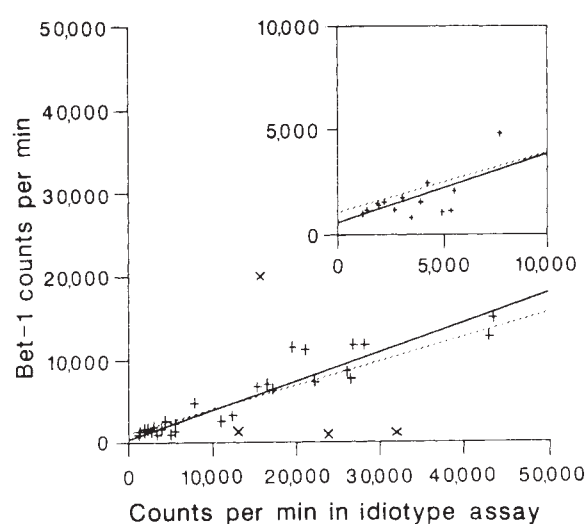
Clones used

As noted above, the authors failed to publish the Bet-1 test results showing that most hybridomas were expressing transgene. The authors did report, however, the results of two tests which they interpreted as showing that most hybridomas were not expressing the transgene. The tests were an S1 nuclease protection assay, and hybridization of northern blots with a 233-bp restriction fragment internal to the V_H gene.

The test results on certain clones were shown by the authors in their Figs 3 and 4. The laboratory records show that according to the authors' own data, the sample of hybridomas chosen by them was not representative with respect to the production of Bet-1 positive material (that is, expression of the transgene). In fact, Table 3 shows that three of five hybridomas shown in Fig. 3 as negative for expression in this assay did not even contain transgene DNA, making these hybridomas completely inappropriate for a test of regulated expression of the transgene. Specifically, . . . those chosen for inclusion in the figures were producing Bet-1 positive material at a lower level than those excluded from Figs 3 and 4. . . Moreover, the data show that both of these RNA tests were in practice less sensitive for the detection of transgene expression than the tests based on direct measurement of transgenic protein. . . But these facts were not reported.

Of the 29 hybridomas in Table 3 said not to express the transgene, the authors reported that nine of these were endogenous idiotype positive. The laboratory records show that every one of these hybridomas was Bet-1 positive as well. The likelihood that this event would have occurred by chance, if transgene and idiotype expression were independent, is very small. Further, the authors selected six endogenous idiotype-positive hybridomas for the sequence data presented in Figs 5, 6 and 7. Every one of these six was also expressing transgene during the process of selection (average of 3,824

FIG.4 Correlation of Bet-1 and idiotype assay. Data from the available laboratory records. All data are shown except for those from the first test on hybridomas from transgenic lymph nodes, which were clearly scrambled (see text). With the omission of the four outlying points (marked 'x'), the data are well correlated. The solid and dotted lines show the least-squares regression lines without and with the outliers respectively. The inset shows the region near the origin.



c.p.m. in the Bet-1 assay). Therefore the data strongly indicate that the selection of idiotype-positive hybridomas was strongly influenced by the expression of μ^a protein.

Discussion

We have analysed laboratory records closely related to a number of central assertions in the paper¹. It is clear from the precise agreement of 14 labelled numerical values that in certain cases the laboratory records contain the data that were actually published. In a number of crucial cases, as noted below, the published paper gives an inaccurate and in fact misleading picture of the underlying data.

(1) A large body of evidence indicates that the Bet-1 reagent, stated by the authors to be a specific anti- μ^a specificity is essential to a central assertion of the paper . . . that certain hybridomas express endogenous μ^b and not exogenous μ^a .

(2) It is clear from the laboratory records that the controls reported in Table 2 were not contemporaneous. The noncontemporaneous controls (1/244) differed by a factor of about 100 from the contemporaneous controls (138/352). The noncontemporaneous controls were reported; the contemporaneous controls were not. The marked dependency of the results of the assays on the length of time in culture indicates that the use of noncontemporaneous controls is not valid.

(3) Hybridomas in the authors' Table 2 were classified as producing exogenous idiotype-positive antibody if they were scored as positive in the idiotype assay and negative in the Bet-1 assay of exogenous μ^a . A central claim of the paper is that most idiotype-positive hybridomas from transgenic mice were producing endogenous antibody. The actual laboratory records of the experiments that were published show this conclusion is wrong. . .

(4) The published paper did not state that despite the threefold difference in sensitivity, both the idiotype and the μ^a assays employed the same cutoff (1,000 c.p.m.). The use of this cutoff was arbitrary, and inspection of

the data shows that it was invalid. The use of this criterion resulted in conclusions not supported by the original data.

(5) The authors published test results in Table 2 on hybridomas from transgenic lymph node despite overwhelming evidence that the test results even though their own tests had proved these results unreliable.

(6) The published paper represents that only 5 of the 34 hybridomas in the authors' Table 3 were expressing transgene. The paper did not report data from protein assays showing that 22 of 34 hybridomas in Table 3 were expressing transgene.

(7) Instead, the paper reported the results of two RNA assays which purported to show that the hybridomas were not expressing transgene. The data not reported show that in practice these assays were less sensitive for the detection of transgene expression than either the authors' μ^a assay or their anti-idiotype assay.

(8) Many clones presented by the authors in Figs 3 and 4 as examples, some described explicitly as "representative", were not in fact typical with respect to expression of the transgene. Some did not even contain transgenic DNA. Therefore the hybridomas shown in Figs 3 and 4 are not appropriate controls to show the absence of transgene expression.

(9) Moreover, of the nine hybridomas from Table 3 stated to be endogenous idiotype positive, every one was expressing transgene at the time it was tested. This is an independent confirmation that expression of idiotype-positive antibody correlates with transgene expression.

The authors' chosen methods of presenting their data convey an inaccurate and misleading picture of their experimental results. The data the authors did not present were directly relevant to the accuracy of their assertions and in fact show many of the most important assertions to be false. Our analysis shows that their experimental results not only failed to support their main conclusions, but in many cases actually contradicted these conclusions.

It is important to note that our criticisms

TABLE 2. Numbers and properties of hybridomas labelled "normal" and "transgenic"

	Normal	Transgenic
Total	352	240
Idiotypic ⁺	4.85×10^5	4.87×10^5
Bet-1 positive	5	42
Bet-1 ⁺ c.p.m.	7.3×10^3	1.17×10^5

Data were obtained from laboratory records. Idiotypic-positive hybridomas were scored by the criteria used in the published paper. Table 1 not shown, data as for Fig. 1.

TABLE 3. Correlation between published DNA data and Bet-1 data from laboratory records

	Data published in Table 3	Bet-1+ Data in laboratory records	Bet-1- Data in laboratory records
DNA+	27	20	7
DNA-	7	2	5

The table shows the number of hybridomas in each category. Hybridomas were scored positive if they gave more than 800 c.p.m. on an anti-idiotypic plate.

do not call into question every part of the published paper we analyse. We do not question the accuracy, for example, of the DNA sequence data presented in the authors' Figs 5, 6 and 7. We have not examined the original laboratory records of the idiotype and other data in Table 1 and Fig. 2, and so we do not have a basis on which to judge these data. However, the data presented in Table 1 and in Figs 2,5,6 and 7 are not relevant to the criticisms presented here.

We have considered the responses of the authors to our analysis, particularly their response that the presentation of their data falls within the norms of standards accepted by the scientific community. The study of this question would require the examination of a number of randomly chosen papers. We do not address this issue, which is in any case irrelevant to the accuracy of the statements made here.

We have also considered the authors' response that it is a matter of interpretation whether or not their paper is accurate. We do not agree. We submit that it is simply a matter of scientific fact that the paper contains serious inaccuracies. We have carefully considered whether data not available to us could significantly alter the conclusions we report here. We conclude that this is most unlikely. Much of our analysis is based on the results of the same experiments that the authors published. We have shown not only that the experiments were misrepresented, but also that the results clearly indicate a conclusion opposite to that stated by the authors. □

Walter W. Stewart and Ned Feder are at the National Institutes of Health, Bethesda, Maryland 20892, USA.

- Weaver, D., Reis, M. H., Albanese, C., Costantini, F., Baltimore, D. & Imanishi-Kari, T. *Cell* **45**, 247-59 (1986).
- Grosschedl, R., Weaver, D., Baltimore, D. & Costantini, F. *Cell* **38**, 647-658 (1984).
- Kung, J.T., Sharrow, S.O., Sieckmann, D.G., Lieberman, R. & Paul, W.E. *J. Immunol.* **127**, 873-876 (1981).
- Velardi, A., Kubagawa, H. & Kearney, J.F. *J. Immunol.* **133**, 2,323-31 (1984).

Imanishi-Kari's riposte

Dr Margot O'Toole's comment (*Nature* **351**, 180; 1991) raised issues about extracts from laboratory notebooks. Dr Imanishi-Kari now replies.

FIRST, I wish to clarify the origin of the "17 pages" of laboratory records, which were copied by Dr O'Toole not from my notebook, but from one kept by Dr Moema Reis, a postdoctoral fellow in my laboratory. They contain unpublished isotype and IgM(μ) allotype immunoassays of supernatants from transgenic cell lines described in Table 3 of the *Cell* paper. They also contain data from a second transgenic mouse that were published as part of Table 2 as well as unpublished data from a "normal" littermate control mouse that we found also to be transgenic.

I also wish to deny Dr O'Toole's account of the meeting with the Wortis committee on 23 May 1986. I did not "candidly admit that certain crucial experiments described in the paper had not even been performed and that the experiments that had been performed had not yielded the results claimed". Accordingly, it is difficult to know exactly what she means by this statement.

Studies

What I know is that, at our 1986 Tufts and MIT meetings, we discussed the following data and experiments:

1. Procedure. In all the immunoassays, anti-idiotypic antibody was used to coat the wells, so that the assays detected only idiotype-positive antibodies. The antibodies in the supernatants that bound to the anti-idiotypic coat were then detected with anti-allotype or anti-heavy-chain isotype or anti-light-chain antibodies.

2. A mistake. Table 2 in the *Cell* paper presents the results of an idiotype and μ allotype (μ^a, μ^b) analysis of hybridoma supernatants from normal and transgenic mice. The data in Table 2 show that only 53 of 172 idiotype-positive hybridoma wells contain secreted IgM. There is a statement in the text that "The remaining 119 clones produced other Ig heavy chain isotypes, the majority being $\gamma 2b$ (data not shown)..."

The phrase in italics is a mistake. The hybridomas in this table were shown to be non-IgM, but were not further isotyped. The supernatants from the hybridoma cell lines shown in Table 3 were isotyped after binding to a plate coated with anti-idiotypic, and many were found to be $\gamma 2b$; this fact was mistakenly put into the text relating to Table 2. We should have, but did not, catch this mistake before the paper was published, but a correction was afterwards published (*Cell* **55**, 541; 1988).

Two things should be noted: (1) this was a mistake, not fraud; and (2) whether the

hybridomas in Table 2 are $\gamma 2b$ is not crucial. What is important is that most of the idiotype-positive hybridomas from the transgenic mice did not make μ^a (the isotype and the allotype of the transgene). Thus the mistake does not alter the conclusions of the paper.

3. The Bet-1 reagent. Two issues have been raised about the use of Bet-1, which is a rat anti- μ^a allotype monoclonal antibody. As we noted in the correction published in *Cell* (reference above), while Bet-1 detects μ^a immunoglobulin (the transgene), it cross-reacts with μ^b (the endogenous IgM) when μ^b is at very high concentrations. So Bet-1 would detect the transgene preferentially, and at much lower immunoglobulin levels than required to detect, by cross-reaction, the endogenous μ^b . But we used Bet-1 to show that transgene-encoded IgM was *absent* from wells that reacted with anti-idiotypic. Cross-reaction with endogenous IgM would *not* have yielded this result. Moreover, the presence or absence of endogenous μ^b was determined by a separate monoclonal anti-allotype antibody. The cross-reaction of Bet-1 is therefore irrelevant and does not affect the conclusions of the paper.

The second issue concerns an experiment found in the "17 pages", where the reagent did not distinguish between μ^a and μ^b . As I have stated before to Dr O'Toole, the data from that experiment were not used in Figure 1 of the *Cell* paper, which is based on another experiment in Dr Reis's notebook. Most of our experiments (including some by Dr O'Toole) gave Bet-1 results similar to those in our Figure 1. Therefore the statement that "experiments that had been performed had not yielded the results claimed" is misleading.

4. The mistyped mouse. Within the "17 pages" there is an immunoassay of hybridoma supernatants from a transgenic mouse and a "normal" littermate control. The "normal" mouse fusion had an unusually high frequency of idiotype-positive antibodies and Bet-1-positive antibodies; it looked like the transgenic mouse fusion.

We knew from our previous work that the idiotype of the μ^a transgene is not detectable in normal mouse serum from C57BL/6 (μ^b) mice (which was corroborated by Dr O'Toole as well as Dr Reis)... we knew that there was something peculiar about the "normal" littermate control. We later analysed the cellular DNA, finding that the mouse was indeed transgenic and had been mistyped as normal.

But the data were not used in the construc-

tion of Table 2, which is based instead on data obtained from fusions of normal C57BL/6 mice. Since the data from the mistyped mouse were not used in the paper, this mistake does not affect the normal control data presented in Table 2 nor the conclusions of the paper. It should be noted that Dr O'Toole was informed after the first Wortis meeting of the mistyping of the mouse, as she acknowledged in her memo to Dr Eisen on 6 June 1986.

On Dr O'Toole's claim that I did not perform the hybridoma "subcloning" (*sic*) data, and that I told her this, I have the following observations:

The primary hybridoma wells that gave rise to Table 2 were cloned by Dr Reis and myself. Dr Reis has informed the Office of Scientific Integrity that the data are to be found in the notebook from which Dr O'Toole copied the "17 pages", but Dr O'Toole has failed to acknowledge this. And she has misconstrued what I told her, which is that *no serological tests* on those wells other than those contained in the 17 pages had been performed.

While it is correct that Dr O'Toole found some mistakes in the *Cell* paper, these have been corrected and do not affect the conclusions. We did not publish any data that seemed to us inexplicable, while the data that were published have been reproduced by us (Dr O'Toole confirmed some as well). My current work confirms the conclusion of the *Cell* paper that mice containing the transgene can generate endogenous B cells

expressing V_H genes that are idiotypically cross-reactive with the transgene idotype and that are generally from the most 3' V_H gene families.

Finally, I wish to restate my objections to the procedures followed by OSI. Its draft report is heavily weighted by the forensic evidence gathered by the Secret Service, but I have been denied access to the experimental details as well as to the original notebooks. I cannot therefore arrange for an independent forensic analysis. Under these circumstances, it would be difficult for anybody to refute a charge of data-fabrication.

For example, I know that the experiments that gave rise to the "June subcloning" and "January fusion" data were performed in 1985. The OSI draft report says that the data appear on "green" tapes, and the Secret Service says it could not find any "green" counter tapes in other people's notebooks after January 1984, whence the conclusion that our data must have been fabricated. I can conclude only that the Secret Service did not look hard enough.

The OSI refuses to tell me whether all of the statements it received were given to the NIH advisory panel before it came to the conclusions in the draft report, and I have just learned that OSI did not send my rebuttal of its draft to the advisory panel. That is why I question the integrity and fairness of OSI's procedures. I am innocent of the charge of fraud, and did not fabricate either the data in the *Cell* paper or the supporting data given to NIH. **Tereza Imanishi-Kari**

existent or fraudulent data". He says my memo is "carefully manicured . . . no hint of fraud" and asks "who misled whom?" But for the five years intervening, Dr Eisen had consistently maintained his inquiry had ruled out the possibility of fraud and misconduct. His 1986 report stated specifically that there was no "evidence of misconduct".

In addition, Dr Eisen's sworn testimony clearly documents that the purpose of his investigation was to find out if fraud was involved. He testified (pages 289, 290) that there was "first this more informal evaluation and then, if there's sufficient grounds . . . we would have gone ahead [with a formal investigation] . . .". In dealing with Dr O'Toole's charges . . . I was aware quite distinctly of the possibility that it [fraud] deserved fair consideration *without her having to charge it*" (emphasis added).

Disavowal

Dr Eisen now disavows his inquiry. He blames its failure on the fact that it began in the absence of an explicit charge of fraud. Certainly my June 1986 memo does not contain the words "fraud" or "misconduct". There are two main reasons for this. First, in 1986, Dr Imanishi-Kari acknowledged openly that the paper contained false claims, and that some of the experiments had not been done. She insisted that these inaccuracies had made their way into the paper as a result of error, not as a result of deliberate fraud. At that time, which was before the large-scale falsification of the laboratory records had occurred, I did not dispute this assertion. As I have testified, I accepted the "error not fraud" explanation. (Much later I discovered that data had been fabricated in an attempt to hide the inaccuracies, and I informed the NIH.)

Second, a MIT administrator told me to remove any suggestion of fraud or misconduct from the memo; she said it was Dr Eisen's job, not mine, to decide whether the actions of my supervisors constituted misconduct. Therefore, my memo did not propose a motive for the misrepresentations.

The laboratory records that I brought to Dr Eisen in June 1986 contained the data for the crucial published experiment, and are accepted as such (1989 hearings, pages 324–334). An analysis shows that all twelve values reported in Table 2 for the transgenic hybridomas had been improperly derived. Dr Eisen now states that it was not clear to him that these records may have represented the only data for one of the tables in the paper. The exact agreement in twelve numbers proves that the notes contained the data for the published experiment.

Besides, saying that he did not understand that there may not have been any other data for the experiment does not excuse the failure of his investigation. By his own admission, he did not see, or even ask to see, any other data for the challenged findings. By Dr Eisen's own sworn account, instead of asking to see the data for the questioned

O'Toole re-challenges

Dr Margot O'Toole replies to Dr Herman Eisen, who investigated the challenge that she made in 1986 to the disputed *Cell* paper.

Dr Eisen writes (*Nature* **351**, 343; 1991) that from his "personal knowledge", he knows my statements to be "inaccurate or grossly to misrepresent the true events". Readers may therefore conclude that it is his word against mine. But that is not the case. Almost everything I say in *Nature* (**351**, 180; 1991) is borne out by subpoenaed documents, forensic evidence, the laboratory notebooks of the authors, data submitted on grant applications, Dr Eisen's written report and his sworn testimony. This documentary evidence stands independently of anything I say, and establishes that it is Dr Eisen's account which is inaccurate.

In his testimony to Congress (House Oversight Subcommittee hearing reports, 4 and 9 May 1989), Dr Eisen said "That whole question of experiments not having been done was not raised by Dr O'Toole [in 1986]". He now again denies that I had indicated in June 1986 that some experiments had not been done, asserting that I made no suggestion that "reported results were based

on nonexistent . . . data". But my June 1986 memo (1989 hearings, pages 307–310) said that some of the experiments had not been done, and this memo was in Dr Eisen's files until it was surrendered to investigators.

Dr Eisen conducted an inquiry into my allegations, and filed a report six months later. He now denies that my challenge had raised the possibility of misconduct. But his report is entitled "Allegations of misconduct . . ." and refers to allegations of misrepresentation brought by me. He wrote that I had raised "a serious question about deliberate misrepresentation of data" (1989 hearings, pages 312–3). He reported that these allegations were unsubstantiated and that the *Cell* paper required no correction. Now that the OSI draft report has said that some of Dr Imanishi-Kari's data are fraudulent, Dr Eisen's statement in *Nature* blames his discredited investigation on me, the witness.

In a complete turnaround, he states that my June 1986 memo contains "no suggestion that reported results were based on non-

experiments, he conducted his inquiry by doing "a lot of talking to people and thinking..." (1989 hearings, page 290).

Now Dr Eisen cites the fact that it took Walter Stewart "weeks" to make sense of my evidence as a defence for the failure of his investigation. First of all, let me point out that Walter Stewart had to start by learning the meaning of "idiotype" and other basic concepts in immunology. It still took him only weeks to conclude that the published table was wrong.

Second, Dr Eisen is a renowned world expert in the relevant speciality, and he had special familiarity with this particular study. He disputes my assertion that a brief examination of the evidence reveals obvious evidence of serious problems. That the evidence establishes that there are obvious problems is documented by the OSI draft report, which refers to it as "*prima facie* evidence" of serious problems (OSI draft report, page 114). Dr Eisen took six months before filing his report and he said in sworn testimony that his investigation "took time" and was "not cursory" (House Oversight Subcommittee hearing report, 12 April 1988, pages 290 and 285). If, during the course of this investigation, he gave only a "brief" examination to the evidence, this was certainly his decision and not mine.

I repeatedly asked Dr Eisen to correct his report and his representations concerning me, but he attacked my character and refused to reconsider (House Oversight Subcommittee hearing report, 12 April 1988, page 99). I wrote a letter to him requesting that he correct the report and help me clear my name. When Dr Eisen's files were surrendered under subpoena, the letter was among them. He had not replied to me. He had merely written sarcastic comments in the margin and filed it. Dr Eisen's statement that I "never asked for [his] help" is thus falsified by documents he himself surrendered under subpoena.

Dr Eisen states that he fears I confuse disagreements with slander and libel. I do not. I maintain that the mere reiteration of the documented facts proves that I do not. Dr Eisen filed a written report stating that I had brought unsubstantiated allegations of misconduct and misrepresentation. For five years, Dr Eisen has vouched for the study in *Cell*; and assured the scientific community that my scientific concerns were trivial.

His description of those concerns was completely inaccurate and as a result I have been subjected to professional ridicule. And his failure to act on my request that he correct his investigative report is still without explanation.

Accuracy

Everything I have said above is part of the record and cannot be denied. The only part of my account not directly shown correct by documentation is my account of the meetings in 1986. My knowledge that certain experiments were not done and that others

were misrepresented came in large part from the 23 May meeting of the "Wortis committee" and the 16 June meeting with Dr Eisen present. Now that the draft OSI report has supported my position, it should be growing clear that the description of how I obtained this detailed and accurate information must also be correct.

Dr Eisen disputes my account of the 16 June meeting, saying he did not hear the admission I heard at the meeting. This denial must be viewed in light of his statement that my memo did not inform him of experiments that were not done; the assertion that he had no reason to consider fraud when his own report stated there was a "serious question about deliberate misrepresentation of data"; and his claim that I never asked for his help and his five years of vouching for the conduct of the authors and the soundness of the study.

Dr Eisen attempts to discredit my account

of the 16 June 1986 meeting by implying that my story has changed over the years. This is not the case. It is easy to verify that the account I give now is the same as that which I gave on the evening of 16 June 1986, during the rest of the summer of 1986, during my interviews with congressional aides and NIH officials in March, May and June 1988 before any of the evidence had been subpoenaed, in my testimony to Congress and throughout the current OSI investigation which began in 1989. (Indeed, a set of documents that appeared to contradict my account later proved to be fabricated.) In fact, as each piece of evidence is uncovered, someone usually has had to change his or her account to make it fit the evidence, but I have never had to do so. That my account has proved consistent with all the evidence that has come to light is no accident. I have been telling the truth all along.

Margot O'Toole

An open letter on OSI's methods

As scientists, we are deeply disturbed by the way in which the charges against Dr Thereza Imanishi-Kari in the well-publicized *Cell* case (leading article in *Nature* 350; 259, 1991) have been handled by the Office of Scientific Integrity of the National Institutes of Health. The need for formal, thorough and fair investigations of possible scientific fraud is clear. However, it is apparent that the procedures followed by the OSI have serious shortcomings, and have not permitted Imanishi-Kari the opportunity to defend herself by a public examination of the evidence against her. Whether or not she is guilty as charged, the precedents which have been set in this case are dangerous. The investigation occurred in a politically charged atmosphere under intense pressure from Congress, which provided the NIH with the Secret Service forensic analysis of Imanishi-Kari's notebooks upon which the serious charges of fabrication are based (News and Comment, *Science* 251; 1552). The confidential draft report of the OSI was broadly leaked to the news media with

devastating impact before Imanishi-Kari could even comment on it, much less cross-examine the witnesses who brought charges against her, or examine and challenge the physical evidence against her. Funding was withdrawn prior to a finding.

The OSI needs to enjoy the confidence of the scientific community, as well as Congress and the public, in order to fulfill its function. The informal and private procedures of the OSI do not provide adequate safeguards for the accused once criminal activity is alleged. The government should not be able to use the OSI procedures to proclaim fraud from behind closed doors, nor should an expert panel of distinguished scientists be made to appear to have acted as a jury under circumstances that did not permit the defence to make a case.

Under the circumstances, we reserve judgement about the facts of this case until Imanishi-Kari has had an adequate opportunity to defend herself. It is not clear to us that the current procedures will allow this to occur. □

SIGNATORIES

Mahmoud Abu-hadid, Jim Allison, Chester Alper, Fred Alt, Ashok R. Amin, Andrei Augustin, Richard B. Bankert, Donald Bellgrau, Michael J. Bevan, Willi K. Born, Margaret Bradley, Patricia Busto, John Cambier, Priscilla A. Campbell, Michael Cancro, Harvey Cantor, Paolo Casali, Jan Cerny, J. Latham Claflin, Edward A. Clark, Nicholas Cohen, Max Cooper, Ronald Corley, Carol Cowing, B. Anne Croy, Chella David, Bernard D. Davis, Dean Dawson, Raymond A. Daynes, Vittorio Defendi, Robert DeMars, Margaret E. Diamond, Martin Dorf, Wesley A. Dunnick, Bo Dupont, Jeannine Durdik, Michael Edidin, Douglas Fallar, Ann Feeney, James Ferrara, John H. Freed, Glen Gaulton, Patricia Gearhart, Victor Ghetie, Laurie Glimcher, Edmond Gold, Daniel Gold, Leonard Herzenberg, Leonore Herzenberg, John A. Hirst, Georg Hallander, Maureen Howard, Leaf Huang, Judith A. Kapp, John Kappler, John Kearney, Raymond Kelleher Jr, Garnett Kelsoe, Norman Klinman, Brian L. Kotzin, Ralph T. Kubo, Pamela J. Lein, Donald Y. M. Leung, Dan Littman, Michael Margolies, Philippa Marrack, Ann

Marshak-Rothstein, Bernard Mathey-Prevot, George L. Mayers, Daniel Meruelo, Richard Miller, John G. Monroe, Claire Moore, Robert Moreland, Donald E. Mosier, Christa Müller-Sieburg, Rebecca O'Brien, Zoltan Ovary, Drew Pardoll, Mary Lou Pardue, David Parker, Jane Parnes, Julia M. Phillips-Quagliata, Carl W. Pierce, Joan Press, Michael Quinones, Klaus Rajewsky, David H. Raulet, Carol Reiss, Elizabeth A. Reparsky, Phillips Robbins, Derry Roopenian, Fred Rosen, Yvonne Rosenstein, Thomas Rothstein, Barry Rouse, David Sacks, Dwayne Savage, Mosello Schnechter, Jonathan Schneck, Dan Schulze, Joel Schwartz, David Scott, Erik Selsing, Jyoti Sen, Ranjan Sen, Eva Severinson, Alan Sher, Hyun Shin, Charles Sidman, Colin Sieff, Philip Siekevitz, Sandra Silberman, Alfred Singer, Kendall A. Smith, Abraham L. Sonenshein, Jonathan Sprent, Uwe D. Steniz, Janet Stavnezer, Lisa Steiner, J. Wayne Streilein, John L. Sullivan, Judy Teale, Jeanette Thorbecke, Annamaria Torriani, Phil Tucker, David Ucker, Martin Weigert, Irving Weissman, Darcy Wilson, Andrew Wright, Sandra J. Yokota, Maurice Zauderer, Ian Zitron, Marjorie B. Zucker

Femtosecond diffraction

John Meurig Thomas

A PROPOSAL to perform electron-diffraction experiments with a resolution of 10^{-15} seconds could revolutionize studies of chemical reactions. Made by Williamson and Zewail earlier this month¹, the proposal suggests that it should be possible to trace the movements of individual atoms in quite complicated molecules as they are rearranged while reactions proceed.

It is a commonplace to remark that as one's ability to identify the intermediates in a chemical or physical process improves so also does one's understanding of the mechanism of the overall change. For this reason biologists concerned with the mode of action of enzymes or hormones, as well as surface scientists whose interests focus on heterogeneous catalysis, or radiation scientists and electrochemists in their diverse fashions, perpetually seek new ways of increasing the temporal resolution of their measurements.

In the course of the past 40 years chemical intermediates of progressively shorter lifetime have been characterized, first on a time-scale of milliseconds then of nanoseconds, with developments in pulse radiolysis, and, with the perfection of lasers and synchrotron sources, of better than picoseconds (10^{-12} s). Impressive as this progress has been in picking up highly populated short-lived species, it is still not good enough to secure an adequate understanding of the mechanism of the critical act of chemical change. Only when one may trace the path of atoms in the course of catastrophic vibrations that sever chemical bonds can the true mechanism of change be properly characterized. This is the essence of transition-state structural chemistry which entails recording events in a time less than it takes a chemical bond to execute a single vibrational cycle — around a picosecond. So Williamson and Zewail's description of an experimental methodology that promises to be able to take diffraction snapshots on the femtosecond (10^{-15} s) time-scale, dramatically increases the prospects of charting the trajectories of individual atoms during the actual turmoil of chemical change.

Since the 22-year-old Lawrence Bragg first used the scattering of X-rays to solve², in 1912, the crystal structure of the mineral zinc blende, X-ray diffraction has remained the single most powerful means of solving molecular structure in atomic detail. Systems of ever-increasing complexity, including enzymes and viruses have, one by one, yielded their structural secrets. But when it comes to determining dynamical changes in molecular entities, X-ray diffraction has been much less revealing than spectroscopic methods. An ostensibly insuperable difficulty with X-ray diffraction is that it requires all the myriad molecules in the crystal to change in synchrony for the resulting diffrac-

tion patterns to chart the individual atomic movements during the instant of chemical change — and that assumes the patterns can be recorded faster than 10^{14} times a second.

In the past few years, however, thanks largely to the endeavours of Zewail and his colleagues in California³⁻⁵, it has proved feasible to expose individual molecules first to an energizing pulse — to spark off chemi-

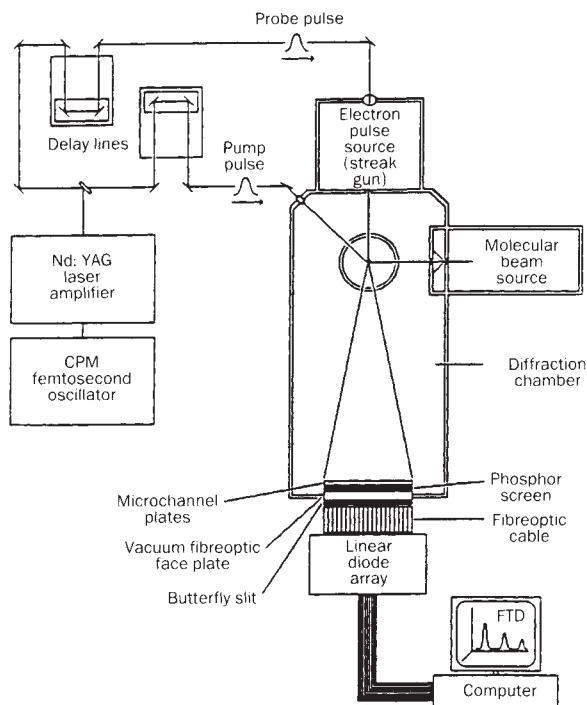


Diagram of the femtosecond transition-state diffraction (FTD) apparatus¹. The pump laser pulse establishes the zero time of the reaction in the molecular beam (of molecules under study). A femtosecond electron burst, generated by the probe pulse, is scattered by the molecular beam, and the resulting diffraction pattern is detected by using a linear diode array. (CPM; colliding-pulse mode-locked.)

cal change synchronously — and then to probe their perturbed structure, spectroscopically, using a second pulse. The duration of the energizing and probe pulses, as well as the time-interval between them (and between repeated probe pulses) is of the order of 10 fs, adequate to explore genuine transition states. In the rather exotic photochemical reaction involving the dissociation of tetrafluorodiiodoethane, $C_2F_4I_2$ into tetrafluoroethene (C_2F_4) and iodine (I_2), real-time studies by femtosecond spectroscopy showed that the reaction proceeds through two consecutive steps, despite the molecular equivalence of the two carbon-iodine bonds. Even more remarkable is the discrepancy in the timescale for bond breakage: the first bond breaks in less than half a picosecond, whereas the second bond takes a hundred times as long to break.

In their study of the bimolecular reaction between HI and CO_2 , Zewail and colleagues

broke the bond between hydrogen and iodine with a pump (laser) pulse. They then observed with probe pulses that the hydrogen attacked the carbon dioxide and stuck to it for hundreds of femtoseconds. The hydrogen atom then stripped one of the oxygen atoms from the carbon: finally, a free hydroxide species emerged five picoseconds after the start of the reaction.

Events of this degree of rapidity cannot be explored at present by X-ray diffraction, notwithstanding the great strides recently made⁶ in synchrotron radiation for time-resolved studies. If, however, monoenergetic beams of electrons were to traverse the perturbed energized species (in the gas phase), the resulting diffraction pattern would yield a snapshot — in fact a Fourier transform of the snapshot — of the real-time structure at the particular instant of the traversal.

What Williamson and Zewail¹ now describe is precisely this extension of femtosecond transition-state spectroscopy to gas-phase electron diffraction. Recent advances made⁷⁻¹⁰ in the design of streak cameras, in which photoelectrons are used to reveal the profile of ultrashort light pulses, convinced Williamson and Zewail that they should be able to generate electron diffraction patterns with femtosecond resolution. If they succeed in so doing, it will be an experimental *tour de force*. As in the previous femtosecond spectroscopic work, two laser pulses are used, the first to trigger the reaction; the second produces a tightly bunched pack of electrons in a modified streak camera, which is then diffracted by the synchronously

evolving molecular beam (see figure). But can the energy and spatial dispersion of the femtosecond 40 keV to 100 keV pulses be rendered sufficiently small to permit the intended measurement? And will the repeat pulses of high-energy electrons be of acceptably (short) duration and constancy of

1. Williamson, J. C. & Zewail, A. H. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5021–5025 (1991).
2. Bragg, W. L. *Proc. Camb. Phil. Soc.* **17**, 43–57 (1912); *Nature* **90**, 402 (1912).
3. Zewail, A. H. & Bernstein, R. B. *Chem. Engng News* **66**, 24–43 (1988).
4. Dantus, M., Bowman, R. M. & Zewail, A. H. *Nature* **343**, 737–739 (1990).
5. Smith, I. W. M. *Nature* **343**, 691–692 (1990).
6. Hajdu, J., Ackery, K. R., Stuart, D. I., Barford, D. & Johnson, L. N. *Trends biochem. Res.* **13**, 104–109 (1988).
7. El-Sayed-Ali, H. E. & Herman, J. W. *Rev. sci. Instrum.* **61**, 1636–1647 (1990).
8. Prokhorov, A. M. & Schelev, M. Ya. *Laser Focus World* **25**, 85–92 (1989).
9. Tebo, A. R. *Laser Focus World* **25**, 21–42 (1989).
10. Niu, H., Degtyareva, V. P., Platonov, V. N., Prokhorov, A. M. & Schelev, M. Ya. *Proc. SPIE Int. Soc. Opt. Eng.* **1032**, 79–85 (1989).

intensity? These are but two key experimental points that cross one's mind.

If the experiment does indeed prove successful — and the authors have decided to investigate the "oscillation" between the covalent NaI and ionic Na⁺I⁻ pair of atoms as their test case — it will mark the dawn of an important new era in structural chemistry, one in which the very act of chemical change is charted in the timescale of the process

MAPPING DISEASE

Seeing flies from space

John Brady

THE tsetse fly, transmitter of trypanosome diseases from game to livestock and humans, is a massive blight on life in Africa. The work of Rogers and Randolph, described on page 739 of this issue¹, by which they demonstrate an effective technique for prediction of areas especially prone to tsetse infestation, introduces a new weapon to Africa's armoury in its attack on the fly.

If the current wars and famines were not enough for the continent to bear, the truly bad news is that Africa's human population of 850 million is increasing in numbers faster than anywhere else in the World; Kenya, for example, will double its present 25 million within the next 20 years. As a result, *per capita* food production is falling inexorably, by 1–2 per cent a year, so more and more people are feeding on less and less. The starting point of agricultural productivity is low anyway: Africa contains 20 per cent of the world's pasture land, yet raises only 10 per cent of the world's cattle, which yield only 3 per cent of the world's meat and milk.

The socio-economic and ecological causes are complex. But in terms of meat and milk production they revolve around the cow, and thus around the tsetse fly. Indeed, by precluding the use of draught animals for ploughing, transport or manure, the trypanosome diseases it carries kept from tropical Africa the agro-economic revolution that occurred in the Middle East around 4000 BC. And the fly exercises the same malign influence today: cattle diseases exclude normal mixed farming from some 10 million km² (an area 20 times the size of France), and human sleeping sickness still debilitates and kills.

Ever since the link between fly and disease was discovered, affected countries have tried to control both. But the results have been scarcely commensurate with the effort — perhaps 5 per cent of the area cleared in 50 years. So the search is on for the means to accelerate the development of effective control measures. With the problems spread on the continental scale, continent-wide solutions need to be found, and this is precisely what Rogers and Randolph have begun to do, with the aid of satellite imagery.

Modern satellites produce data that can be analysed to reveal the ecological condition of

itself. The authors have already simulated experimental radial distribution curves for excited NaI as a function of time. What we now await is the actual experimental determination. □

John Meurig Thomas is at the Davy Faraday Research Laboratory, Royal Institution, 21 Albemarle Street, London W1X 4BS, and the University of Wales, University Registry, Cathays Park, Cardiff CF1 3NS, UK.

the lands they pass over. Especially useful is the normalized difference vegetation index (NDVI) which, in effect, indicates the level of photosynthetic activity, and thus the vegetation cover². From this can be inferred the levels of recent rainfall. Monthly NDVI images covering the whole of Africa are available³, and they reflect in a stunningly impressive way the moving seasonal rainfall as it sweeps to and fro across the continent.

These data have obvious potential for warning of ecological change such as drought⁴, but have also begun to be analysed for what they may indicate of the distribution of animals, through correlations between habitat characteristics and vegetation indices. For instance, on the macro, country-wide scale likely locust breeding zones in Mali have been found; and on the micro scale, individual Landsat pixels have been used to pick out the feeding grounds of waders in the threatened moorlands of northern Scotland. What Rogers and Randolph have done is to use NDVIs to reveal areas of maximum tsetse infestation, aiming to do this soon for the whole continent.

The potential for exploiting NDVIs in disease prediction had been hinted at in work that identified, after the event, periods and areas in Kenya likely to be subject to outbreaks of the mosquito-borne cattle disease, East Coast fever⁵, but this — as with the locusts in Mali — required satellite data concurrent with the outbreaks being predicted. Rogers and Randolph, however, have taken a synoptic view, and reveal that NDVIs averaged over several years in the 1980s correlate beautifully with long-term surveys of tsetse flies made in the 1970s and much earlier. Moreover, they have correlated NDVIs with habitat characteristics and animal populations in a rigorous, statistical sense. The authors show not only how the seasonally changing NDVIs across the year relate to each following month's tsetse population (because breeding success depends partly on saturation deficit, which is associated with rainfall), but also how annually averaged NDVIs indicate where — geographically — tsetse survival rate will be highest, and hence the risk of trypanosome transmission greatest.

Bearing in mind that the area potentially

infested with the fly is covered by only a few hundred tsetse entomologists (say 50,000 km² per entomologist?), the gaps in our knowledge of tsetse distribution and the consequent trypanosome 'challenge' are — like everything else in the continent — on an epic scale. Rogers and Randolph may have taken a key step in crossing some of those gaps, in that their technique should pin-point areas for control.

For example, there is a large belt of tsetse flies running through the countries south of Malawi. By 1985 Zimbabwe was on the way to ridding itself of its own share of the flies and was looking for the means to prevent re-invasion from its neighbours. It approached the European Community for aid to solve the problem by attacking the whole belt. The EEC was on the point of agreeing to finance aerial spraying of the entire 35,000 km², when a now infamous full-page article by the late Marcus Linear in *The Sunday Times* in effect killed the programme, by influencing the Council of Ministers against the use of broadcast insecticide (even though treatment would have been at less than 10 g ha⁻¹).

Although one may question the right of the West's eco-lobby to interfere in the Third World's solutions to its own desperate problems, there was one happy result from the ensuing delay: a more environmentally friendly technology caught up with the aeroplanes, and simple traps will now be used instead of aerial spraying.

These traps (or, more properly targets) were developed in Zimbabwe by G. A. Vale and his colleagues. They consist of a rectangle of black cloth on a wire frame baited with key components of cow odour and doped with a quick-acting, pyrethroid insecticide. Tsetse flies are irresistibly attracted upwind by the smell, see the black target and — perhaps taking it for a cow — fatally land on it. So effective are the targets that by using them alone (at a mere 4 km⁻²) Vale succeeded within 6 months in exterminating all tsetse from a 600 km² trial area in Zimbabwe⁶.

So now the EEC programme will use targets. Compared to the bulldozing technique of aerial spraying, this is like using a stiletto, and demands the stiletto's precise handling: targets work well only when sited with great care. Hence Rogers and Randolph's method of identifying high infestation zones will provide an important way of helping to optimize the distribution of targets over the vast areas in which it is now planned to site them. □

John Brady is in the Department of Biology, Imperial College, Silwood Park, Ascot, Berkshire SL5 7PY, UK.

1. Rogers, D. J. & Randolph, S. E. *Nature* **351**, 739–741 (1991).
2. Sellers, P. J. *Int. J. Rem. Sens.* **6**, 1335–1372 (1985).
3. Tucker, C. J., Townshend, J. R. G. & Goff, T. E. *Science* **227**, 369–375 (1985).
4. Henriksen, B. L. & Durkin, J. W. *Int. J. Rem. Sens.* **7**, 1583–1608 (1986).
5. Linthicum, K. J., Bailey, C. L., Davies, F. G. & Tucker, C. J. *Science* **235**, 1656–1659 (1987).
6. Vale, G. A., Lovemore, D. F., Flint, S. & Cockbill, G. F. *Bull. ent. Res.* **78**, 31–49 (1988).

Hue and the heptahelicals

John Mollon

MUCH of the intercourse between our nerve cells, and much of our interaction with the outside world, depends on members of the same superfamily of protein molecules¹. These molecules, embedded in cell membranes, are characterized by seven helices that span the membrane, and their action within the cell is achieved by G proteins. Commonly (and clumsily) known as G protein-coupled receptors, they might be better named heptahelical receptors. Already the genes for several hundred heptahelicals have been sequenced; R. J. Lefkowitz provided a helpful discussion of them and summarized the known ligands in *News and Views* last month².

The opsins (the protein parts of the photosensitive pigments of the retina) are particularly well-studied members of the heptahelical family. Because they exhibit a rich polymorphism in humans and primates, and because a number of pathological mutations have been identified (in cases of retinitis pigmentosa^{3,4}), they offer useful models to those who study their more distant relatives⁵. Now, in *Science*⁶, Neitz *et al.* put forward a theory of how the middle-wave and long-wave cone pigments of human vision come to differ in their spectral absorbances (that is, in the way their sensitivity varies with the wavelength of the incident light).

Since 1986, when Nathans and his colleagues published sequences for the human cone pigments, it has been known that the long-wave and middle-wave opsins differ only in 15 amino acids and only seven of these are non-homologous substitutions lying within the transmembrane regions where they are likely to affect the absorbance properties of the molecule⁷. To home in on the amino-acid sites that are critical, Neitz *et al.* have now exploited the striking polymorphism of colour vision found in many New World monkeys, such as marmosets, tamarins and squirrel monkeys⁸. Within a given species, the hue discrimination of the male monkeys resembles that of dichromatic men (that is, men who are colour blind in the sense that they can match all colours with two variables), whereas most of the female monkeys are trichromatic, enjoying good discrimination within the red-green range. Moreover, within each species, there appear to be at least three kinds of dichromat and three kinds of trichromat, differing in their exact spectral sensitivity.

A decade ago, Bowmaker, Jacobs and I showed that these behavioural variations are associated with variations in the retinal photopigments⁹. That work led to a genetic model which supposes that New World monkeys, unlike ourselves, have only a single locus for specifying a pigment with peak sensitivity in the red-green range of the spec-

trum¹⁰. This locus, like the loci for the human long- and middle-wave pigments, is thought to be on the X chromosome. Three alleles can occur at the monkey's single locus, so giving the three kinds of male dichromat. The three alleles occur with similar frequency, so that many females will be heterozygous, inheriting different versions of the gene on their two X chromosomes. Owing to X-chromosome inactivation, the heterozygous female will express only one allele in any given cell; and apparently her visual system can exploit the presence of two subtypes of cone. As all the monkeys also have a short-wave pigment, the heterozygote becomes trichromatic.

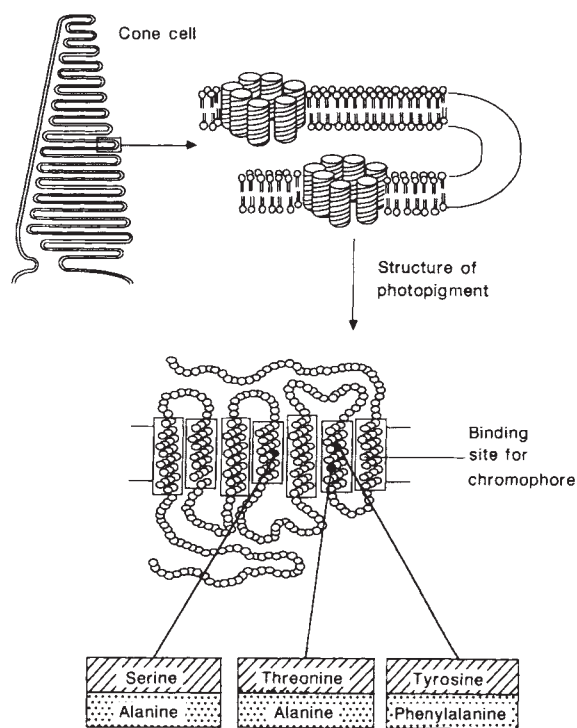
These polymorphisms offer an attractive way of discovering the molecular differences that give a photopigment its spectral sensitivity. Neitz *et al.* have directly sequenced the genes for the 562-nm, 556-nm and 541-nm pigments that occur individually in different male tamarins, and the 561-nm, 547-nm and 532-nm pigments that occur in male squirrel monkeys. They find pairs of pigments that differ at only one of the seven sites considered relevant in the case of human long-wave and middle-wave pigments. Thus the 556-nm tamarin pigment differs from the 562-nm tamarin pigment only by the substitution of alanine for serine at position 180; the 541-nm tamarin pigment differs from the 556-nm tamarin pigment only by the substitution of alanine for tyrosine at position 285; and the 547-nm squirrel monkey pigment differs from the 556-nm tamarin pigment only by the substitution of phenylalanine for tyrosine at position 277 (see figure). In each case a non-polar amino acid replaces a hydroxyl-bearing residue.

Neitz *et al.* are led to the bewitchingly simple theory that the three sites (180, 277 and 285) account for all the variation in the spectral sensitivity of primate cone pigments in the range 530 to 565 nm and that the three substitutions are linearly additive in the number of nanometres by which they shift the peak sensitivity of the photopigment. The best estimate for the shift produced by the substitution at site 180 is 5.3 nm; and for the substitutions at sites 277 and 285 the best estimates are 9.5 and 15.5 nm respectively. These three

substitutions sum to give the approximately 30 nm difference between the human long- and middle-wave pigments.

How secure is this additive theory? Neitz *et al.* show that it accurately predicts paired comparisons among all eight of the pigments considered and they also show that the substitutions observed at other sites cannot predict such additive shifts. But they restrict themselves to the subset of alternative theories in which individual substitutions act alone in an additive way; they do not consider more complicated theories in which additional, facilitating, substitutions are required to produce the full 30-nm shift between the human long-wave and middle-wave pigments. The theory can be tested in two ways: one is to exploit nature's own experiments and to sequence more opsin genes; the other is to change individual amino acids by site-directed mutagenesis and measure the expressed pigment — as has already been done with rhodopsin^{11,12}.

Neitz *et al.* end their paper by suggesting that two forms of the long-wave opsin are present in the human population (a possibility earlier suspected from microspectrophotometric results¹³): they propose that the two pigments differ according to whether alanine or serine is present at site 180 and



Visual transduction depends on a G protein-coupled molecule embedded in the multiply infolded membrane of the outer segment of a cone cell. The opsin (the protein part of the molecule) is bound to the chromophore, 11-*cis*-retinal, at the site indicated in the lower part of the diagram (seventh helix). The additive theory of Neitz *et al.* proposes that just three amino acids (filled circles in the cartoon of the molecule) determine the spectral difference between the human long- and middle-wave pigments. The pairs of alternative amino acids are shown in boxes at the bottom of the diagram and in each case it is the upper of the two that produces a red shift in the absorbance of the molecule. (Figure by P. Jeffs.)

RÉSUMÉ

thus differ by 5 to 6 nm in the wavelength of maximum absorbance. Earlier, in 1986, Neitz and Jacobs suggested that men fall into two clear groups on the basis of their subjective colour matches¹⁴. The finding has been controversial¹⁵, and at the recent meeting of the Association for Research in Vision and Ophthalmology, J. Neitz implicitly withdrew the claim¹⁶. What he and his colleagues now suggest is that a subset of men carry copies of two forms of the long-wave gene, expressing them in different cones. Such 'pseudo-heterozygotes' may reveal themselves by changes

in their colour matches after selective colour adaptation. It was in these pages 110 years ago that Lord Rayleigh suggested that different observers may live in slightly different perceptual worlds¹⁷, and some of this variance can now be traced to variations in heptahelical receptors. Perhaps polymorphisms of other heptahelicals account for some of the variation that is observed in people's mental worlds. □

John Mollon is in the Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

Rat hole

SHIP rats (*Rattus rattus*) were curiously absent from the city of York in the Dark Ages (fifth–eighth centuries AD), even though they were common in Roman times and again during the ninth-century Viking period. "Absence of evidence is never conclusive evidence of absence" says T. P. O'Connor (*J. Zool.* **224**, 318–320; 1991), "though in this case it is quite compelling". Analysis of an eighth-century deposit revealed abundant bones of mice (*Mus spp.*) but no rat material whatsoever, an observation that carries statistical weight because of the conspicuous presence of rats at other times. It seems that rats introduced by the Romans subsequently disappeared, though no Pied Piper has yet been identified.

Inhibited hopes

TERGEMININ, echistatin, eristicophin, barbourin — these resonant names attach to a series of snake venom proteins, all of them antagonists of the integrin family of cell-surface receptors and known accordingly as disintegrins. Their activity centres on three residues, which are also the sequence that the integrins recognize in adhesion proteins. The venoms inhibit platelet aggregation, by binding to a glycoprotein complex, but the therapeutic potential of a disintegrin of broad specificity is limited because of the mayhem it would cause elsewhere. Now, R. M. Scarborough *et al.* (*J. Biol. Chem.* **266**, 9359–9362; 1991) have brought to light barbourin (courtesy of the southeastern pigmy rattlesnake). In barbourin the inhibitory sequence Arg-Gly-Asp has become Lys-Gly-Asp, and this is highly specific for the platelet receptor — which should commend it to clotters everywhere.

Standard deviation

ARE the first cracks showing in the standard model of particle physics? Workers on the ALEPH experiment at CERN's Large Electron-Positron collider reported data to a CERN seminar last week that could suggest so. Out of 200,000 Z^0 particles detected at ALEPH, 35 decayed by an unusual route to produce a pair of leptons (an electron, muon or tau and its respective antiparticle) and a short-lived photon. Of these few decays (actually an unexpectedly large total), more resulted in tau particles (15) than electrons or muons (10 each), where the reverse pattern was expected. With such a small sample, statistical fluctuations could be laying a false trail, but ALEPH workers say there is only a 1 in 100 chance this is so. With further data being analysed, confirmation or refutation may soon be at hand.

- Hall, Z. *Trends Neurosci.* **10**, 99–101 (1987).
- Lefkowitz, R. J. *Nature* **351**, 353–354 (1991).
- Dryja, T. D. *et al. Nature* **343**, 364–366 (1990).
- Bhattacharya, S. *et al. Lancet* **337**, 185 (1991).
- Mollon, J. D. & Jordan, G. *Die Farbe* **35/36**, 139–170 (1989).
- Neitz, M., Neitz, J. & Jacobs, G. H. *Science* **252**, 971–974 (1991).
- Nathans, J., Thomas, D. & Hogness, D. S. *Science* **232**, 193–202 (1986).
- Jacobs, G. H. *Vision Res.* **23**, 239–248 (1983).
- Jacobs, G. H., Bowmaker, J. K. & Mollon, J. D. *Nature* **292**, 541–543 (1981).

- Mollon, J. D., Bowmaker, J. K. & Jacobs, G. H. *Proc. R. Soc. Lond. B* **222**, 373–399 (1984).
- Sakmar, T. P., Franke, R. R. & Khorana, H. G. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8309–8313 (1989).
- Nathans, J. *Biochemistry* **29**, 937–942 (1990).
- Dartnall, H. J. A., Bowmaker, J. K. & Mollon, J. D. *Proc. R. Soc. Lond. B* **220**, 115–130 (1983).
- Neitz, J. & Jacobs, G. H. *Nature* **323**, 623–625 (1986).
- Jordan, G. & Mollon, J. D. *Investig. Ophthalmol. vis. Sci.* **29**, suppl., 164 (1988).
- Neitz, J., Neitz, M. & Jacobs, G. H. *Investig. Ophthalmol. vis. Sci.* **32**, 1092 (1991).
- Rayleigh, Lord *Nature* **25**, 64–66 (1881).

COSMOLOGY

A beast in the menagerie

Simon Lilly

If bright astronomical objects are telling us that something special is going on elsewhere in the Universe, then the exceptionally luminous object that Rowan-Robinson *et al.* describe on page 719 of this issue¹ is well worth noting. Designated 10214+4724, the object has a redshift of 2.286, indicating that it is among the more distant known objects and that we are seeing it as it was when the Universe was between a sixth and a third of its present age. It was found through the systematic identification of over a thousand sources catalogued in the Faint Source Survey of the Infrared Astronomy Satellite (IRAS). Its infrared luminosity, which exceeds a hundredfold its optical luminosity, is equivalent to that of 3×10^{14} Suns and exceeds by a small margin, the authors say, that of any other known object.

The discovery of such an object immediately raises the question of its cosmological significance. Where does this object fit in to the menagerie of high-redshift objects and what if anything does it tell us about the formation of galaxies and galaxy clusters in the early Universe?

There are two unresolved uncertainties concerning the object that need to be addressed before these questions can be answered definitively. First, where does all the infrared radiation come from? From dust surrounding and heated by a powerful primary source seems the most likely answer, although the infrared could be synchrotron emission from a highly self-absorbed, compact source of relativistic electrons. Second, if dust is responsible, what is the primary source? It could be a quasar-like galactic nucleus, in which case it must be as luminous as any known. Or the dust could be shroud-

ing an outburst of star formation of enormous proportions, in which case the equivalent of 10^5 new Suns are forming every year, enough to make a very large galaxy in only 10^7 years.

Whatever generates the primary radiation, if dust is the source of the secondary, infrared radiation, as seems likely, then 10214+4724 contains an impressive amount of it. The mass can be estimated from the thermal spectrum if assumptions are made about the grain size and the spatial distribution of the dust, which affects its temperature distribution. (Unfortunately, IRAS could not see the coolest dust.) Making various plausible assumptions, Rowan-Robinson *et al.* estimate there are $4 \times 10^8 - 10^{10}$ solar masses of dust, although this range could no doubt be extended both ways. Our Galaxy, by comparison, contains only 2×10^8 solar masses of dust; its total metal content, comprising elements synthesized in stellar burning, is roughly 4×10^9 solar masses; this resides mostly in formed stars. So, according to Rowan-Robinson *et al.*, a significant fraction of the metals that we see in a typical galaxy now had already been produced by the time of 10214+4724, perhaps 2.3 billion years after the Big Bang.

There are other indications that gas in the Universe was significantly enriched with metals at these high redshifts. For instance, the emission-line spectra of quasars at redshift $z > 4$ indicate normal relative abundances² (although the abundance relative to hydrogen is poorly determined). Distant galaxies associated with powerful radio sources seem already to contain the bulk of the (enriched) stellar population of a distant giant elliptical galaxy³. And the cos-

mological clouds responsible for damped Lyman- α absorption lines in quasar spectra cause reddening of the background quasar that indicates a dust/gas ratio that is 0.02–0.25 times that in the present-day Milky Way⁴. The large quantity of dust inferred to be in 10214+4724 thus supports the notion that, for some objects at least, a substantial fraction of their present-day stars and metal abundances were in place by $z = 2.5$.

Unfortunately, 10214+4724 and objects like it are extremely rare even at $z = 2$, so the argument of early formation may apply only to a small fraction of galaxies. Rowan-Robinson *et al.* estimate that there may be only one such object in every 10^9 cubic megaparsecs, which is roughly the same as for ultraluminous optical quasars at $z = 2.5$ (which, interestingly, would have about the same total luminosity as 10214+4724) and for powerful radio galaxies at $z > 3$. By comparison, the present number density of typical galaxies is a million times greater and even very massive galaxies are 1,000 times more common.

But, as with other classes of transient objects, in estimating the number density of the parent population, one must allow for the short life of the high-luminosity phase. Given that the luminous phase of 10214+4724 could last as little as 10^7 years, the background dormant population could be 100 times greater. Furthermore, less extreme versions of 10214+4724 may be lurking with lower luminosities.

Deep spectroscopic surveys of optically-selected faint objects have been carried out⁵ which would certainly have identified similar objects up to three times fainter in the optical waveband. But the number studied is too small for there to be any serious constraints on the density of subdued relatives of 10214+4724. An interesting approach is to look at the radiation density produced by the population which, from the observations of Rowan-Robinson *et al.*, must be around 3×10^{48} J Mpc⁻³. This is a thousand times less than the radiation density produced by the flat-spectrum blue galaxies seen⁶ at faint optical magnitudes (around 3×10^{51} J Mpc⁻³). So, even if the luminosity in this class of object comes from newly formed stars, it represents only a small part of the total star-formation in the Universe.

My own best bet for 10214+4724 is that it is an obscured quasar of very high luminosity. This naturally accounts for the high-excitation emission lines seen in the optical spectrum and removes the need for what is by any account an uncomfortably high star-formation rate. It follows the interpretation⁷

placed on what is probably the closest analogue to 10214+4724 at lower redshift, IRAS 09104+4109 at $z = 0.442$. Also, the fact that the implied number of such objects at high redshift is comparable to that of optically selected quasars of equivalent total luminosity could reflect the approximate equality at low redshift between the number of infrared-bright and conventional optically bright quasars. The presence of large amounts of dust in quasars at high redshift (with the implied early processing of material through stars) then ties in with what we know of other objects at these high redshifts.

AIDS

The positive effect of the negative factor

Bryan R. Cullen

ALTHOUGH Nef is the largest of the auxiliary gene products encoded by human immunodeficiency virus type 1 (HIV-1), its negligible effect on the level of viral replication in culture has made it the most enigmatic. Writing in *Cell*, Kestler *et al.*¹ now reveal that a functional *nef* gene is essential for efficient viral replication and progression to AIDS in rhesus monkeys infected with the related simian immunodeficiency virus (SIV).

The *nef* gene is observed in all primate lentiviruses, but it is highly polymorphic and also obviously defective in many replication-competent proviral clones. Early attempts to define the phenotype of its protein product revealed little effect on either HIV-1 replication or cytopathicity^{2,3}. These studies did however report a slight but apparently significant reduction in the efficiency with which *nef* expressing HIV-1 was able to spread in culture, leading to the perhaps premature designation 'negative factor' (Nef). Subsequent reports suggested that the cytoplasmic Nef protein might act as a modest but specific repressor of HIV-1 transcription^{4,5}, and even proposed that it is involved in the establishment of viral latency *in vivo*^{5,6}. Yet several other studies using identical *nef* isolates were unable to detect any negative effect on HIV-1 transcription or replication^{7,8}, thus implying that a clear understanding of Nef's participation in the HIV-1 replication cycle was unlikely to emerge solely from investigations of viral replication in culture.

It is perhaps this consideration that prompted Kestler *et al.*¹ to compare the effect of Nef expression on the replication of SIV *in vitro* with that observed in an animal host, the rhesus monkey. The SIV proviral clones analysed differed only in that they contained a wild-type *nef* gene (Nef open), a *nef* gene bearing a premature translation termination codon (Nef closed) or a deletion in *nef* (Nef deleted). As implied by some of the earlier studies in HIV-1, these three otherwise isogenic viruses displayed indistin-

Whatever the explanation of 10214+4724's extreme luminosity, it provides a timely reminder that there are many classes of object at high redshifts that may be cosmologically significant, but are relatively invisible to conventional astronomical techniques. As we seek to make out the dim and murky form of the cosmological dog at high redshift, we must be sure we are not too distracted by the wagging of the optical tail. □

Simon Lilly is in the Department of Astronomy, University of Toronto, Toronto, Ontario M5S 1A7, Canada.

guishable replication patterns in primary rhesus CD4⁺ T cells and macrophages, and in transformed T-cell lines in culture. But a dramatic difference in replication potential *in vivo* soon became evident. In particular, whereas 'Nef open' SIV replicated to high titres in infected monkeys, 'Nef deleted' SIV gave rise to barely detectable levels of virus replication, even though a robust immune response demonstrated that infection had occurred. Subsequently, monkeys infected with the cloned 'Nef open' SIV showed a marked decline in the number of circulating CD4⁺ lymphocytes and developed an overt immunodeficiency. In contrast, the monkeys infected with the attenuated 'Nef deleted' virus remained healthy and retained a normal CD4⁺ T-cell count.

Perhaps the most striking observation resulted from analysis of monkeys infected with the 'Nef closed' SIV. Like the animals infected with 'Nef open' SIV, they displayed readily detectable SIV titres and went on to develop AIDS. Remarkably, analysis of virus recovered at various stages after infection revealed that the *nef* gene had invariably reopened due to any one of a series of different mutations that specifically removed the inserted translation termination codon — even as early as two weeks after infection, four out of four *nef* genes recovered from a rhesus monkey infected with the 'Nef closed' virus revealed uninterrupted *nef* open reading frames.

These results allow two conclusions to be drawn. First, a functional SIV *nef* gene product is essential for efficient viral replication and pathogenesis in infected rhesus monkeys — Nef is clearly a positive factor *in vivo*. Second, the critical importance of Nef for maintenance of SIV replication *in vivo* is not reproduced by current assays of virus replication *in vitro*.

One possible reason for this failure, as noted by Kestler *et al.*¹, is that *in vitro* assays do not reproduce many of the normal host responses to viral infection. Host inflam-

1. Rowan-Robinson, M. *et al.* *Nature* **351**, 719–721 (1991).
2. Schneider, D. P., Schmidt, M. & Gunn, J. E. *Astr. J.* **98**, 1507–1522 (1989).
3. Lilly, S. J. in *The Evolution of the Universe of Galaxies* (ed. Kron, R. G.) 344–354 (Astr. Soc. Pac., San Francisco, 1990).
4. Fall, S. M., Pei, Y. C. & McMahon, R. G. *Astrophys. J.* **351**, L5–L8 (1989).
5. Colless, M. N., Ellis, R. S. & Shanks, T. *Mon. Not. R. astr. Soc.* **244**, 408–423 (1990).
6. Tyson, J. A. *Astr. J.* **96**, 1–23 (1988).
7. Kleinmann, S. G. *et al.* *Astrophys. J.* **328**, 161 (1988).

matory and immunological reactions can strongly inhibit viral replication, and there are several examples of viral gene products that are specifically designed to prevent antigen presentation by infected cells or to inhibit cytokine function. But the observation that lack of *nef* function dramatically inhibits the rate of SIV replication within the first two weeks post-infection — that is, before the host immune response is fully established — may mean instead that a primary effect at the level of viral gene expression is involved. No explanation of why current *in vitro* assays for SIV and HIV-1 replication would fail to reproduce this second effect has been proposed.

A molecular explanation for the action *in vivo* of Nef is unlikely to be forthcoming until this phenomenon can be reproduced in culture. The little that is known about the properties of the Nef protein, while intriguing, does not suggest an obvious solution. Both the HIV-1 and SIV Nef proteins are myristylated at their N termini and, presumably as a result, are predominantly associated with cytoplasmic membrane structures⁸. This would clearly be an excellent location from which to affect either protein transport or membrane signalling. It has, in fact, been proposed that Nef can down-regulate cell-surface expression of CD4, the receptor for both HIV-1 and SIV, through a post-translational mechanism^{9,10}. But it is unclear how this somewhat controversial⁶ observation could explain the immediate and intense selection for Nef function observed *in vivo*. It has also been suggested that Nef can function as a guanine-regulatory (G-) protein¹⁰, raising the possibility that Nef might act by interfering with the normal function of an undefined G-protein-linked cellular regulatory pathway. This hypothesis also remains unconfirmed.

Although the role of the Nef protein in the HIV-1 replication cycle therefore remains obscure, it is known that Nef is an 'early' gene product¹¹, a distinction it shares with the essential nuclear regulatory proteins Tat and Rev. Remarkably, up to 80 per cent of the early, multiply spliced class of viral transcripts encode Nef¹². Later in the viral replication cycle, expression of Nef messenger RNA is down-regulated by the action of the viral Rev *trans*-activator¹¹. So Nef, like Tat and Rev, most probably

functions soon after infection.

The data of Kestler *et al.*¹ indicate that Nef is equally critical for the establishment and maintenance of a pathogenic SIV infection *in vivo*. If this SIV model accurately reflects the situation in HIV-1, then an obvious approach to the development of a live, attenuated vaccine strain of HIV-1 presents

MANTLE PLUMES

Smoke signals from the deep

Laura Garwin

THE Earth's drifting continents bear witness to its active interior. The 'conveyor belts' of ocean crust, cooling as they travel from mid-ocean ridge to subduction zone, release about 85 per cent of the Earth's heat flow. But this most obvious surface expression of the convecting mantle reveals less than might be expected of the pattern of flow beneath: the mantle material that rises beneath ridges seems to do so passively, to fill the gap left by the spreading plates, rather than marking the location of the upwelling limb of a mantle convection cell. With flagrant disregard for symmetry, the characteristic upwellings of mantle convection seem not to be sheet-like counterparts of the slabs sinking at subduction zones, but cylindrical 'plumes' whose relationship to the rigid plates overhead is at

itself. It will clearly be of great interest to determine whether monkeys infected with the 'Nef deleted' SIV are protected against challenge with pathogenic SIV isolates. □

Bryan R. Cullen is in the Howard Hughes Medical Institute and Section of Genetics, Duke University Medical Center, Durham, North Carolina 27710, USA.

best indirect. But as a recent meeting* showed, although these plumes transport only about 10 per cent of the Earth's escaping heat, they carry much valuable information about the pattern of flow in the mantle.

Mantle plumes (Fig. 1) were first proposed as an explanation for 'hotspots' — linear chains of volcanoes that seem to arise from a stationary smokestack in the deep mantle, which releases the occasional puff into the drifting plate overhead. Wilson¹ imagined a Hawaiian plume rising from the relatively static core of a mantle convection cell, whereas Morgan² suggested that plumes tap relatively primordial material from the lower mantle. The latter idea rapidly gained support from geochemistry, as the ocean island basalts formed at hotspots sample a reservoir that is more enriched in the melt-loving 'incompatible' trace elements than the mantle source of mid-ocean-ridge basalts. Moreover, because the enrichment applies to elements, such as strontium and samarium, that are daughters or parents in long-lived radioactive decay schemes, one can infer the time over which the different reservoirs have been isolated from one another: estimates range from 500 to 2,000 million years — a good part of the Earth's age. The 'primordial' nature of the plumes seemed to be reinforced by the presence of large amounts of the isotope ³He in some hotspot basalts; this isotope is neither produced by natural radioactive decay nor recycled from the surface back into the Earth's interior, and hence must have survived from the time of the Earth's formation.

These observations and others led to a tacit acceptance by most geochemists that the mantle convects in two layers, separated by the well-known seismic discontinuity at 670 km depth. Mid-ocean ridges, so the story went, sample the depleted upper mantle (depleted by the extraction of the continental crust early in Earth history, and to a lesser extent by the extraction of mid-ocean-ridge basalt itself). Ocean islands represent material carried by plumes from the lower, 'primitive' mantle, which does not participate in the plate-tectonic cycle, and thus retains its primordial composition.

But the evidence for the existence of primitive mantle has been crumbling, and

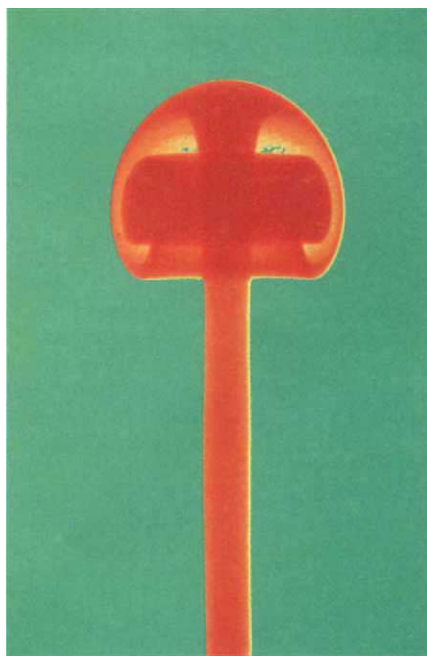


FIG. 1 Laboratory model of a mantle plume, created by pumping hot, dyed syrup (red) into the bottom of a deep tank. As the plume ascends, a large volume of surrounding syrup (yellow) has been heated and stirred into the plume head. By this process of entrainment, plume heads from the core-mantle boundary can increase their volume by up to 50 times during their ascent through the mantle. (Courtesy of R. Griffiths, M. Fitzpatrick and R. Wyde-Browne. Image enhanced by false colour.)

1. Kestler, H. W. III *et al. Cell* **65**, 651–662 (1991).
2. Terwilliger, E., Sodroski, J. G., Rosen, C. A. & Haseltine, W. A. *J. Virol.* **60**, 754–760 (1986).
3. Luciw, P. A., Cheng-Mayer, C. & Levy, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1434–1438 (1987).
4. Ahmad, N. & Venkatesan, S. *Science* **241**, 1481 (1988).
5. Niederman, T. M. J., Thielan, B. J. & Ratner, L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1128–1132 (1989).
6. Cheng-Mayer, G., Iannello, P., Shaw, K., Luciw, P. A. & Levy, J. A. *Science* **246**, 1629–1632 (1989).
7. Kim, S., Ikeuchi, K., Byrn, R., Groopman, J. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9544–9548 (1989).
8. Hammes, S. R., Dixon, E. P., Malim, M. H., Cullen, B. R. & Greene, W. C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9549–9553 (1989).
9. Garcia, J. V. & Miller, A. D. *Nature* **350**, 508–511 (1991).
10. Guy, B. *et al. Nature* **330**, 266–269 (1987).
11. Malim, M. H., Hauber, J., Fenrick, R. & Cullen, B. R. *Nature* **335**, 181–183 (1988).
12. Robert-Guroff, M. *et al. J. Virol.* **64**, 3391–3398 (1990).

* Caltech Mantle Plume Symposium, Pasadena, 2–4 May 1991.



FIG. 2 Results of a two-dimensional numerical convection experiment, modelling the formation of ocean crust at a spreading centre and its subsequent recycling into the mantle at a subduction zone. The different shades correspond to composition: white is pure ocean crust; black is the mantle residue after crust extraction; grey is 'fertile' mantle that has not had crust extracted. About a tenth of the subducted crust segregates to the bottom of the system, which could be either the core-mantle boundary or the bottom of the upper mantle. The time scales with the depth of the layer: for whole-mantle convection, subducted crust can be stored at the base of the mantle for 1,000 million years or longer. (Courtesy of U. Christensen.)

the idea had few advocates at the meeting. Al Hofmann (Max-Planck-Institut für Chemie, Mainz), speaking of "the myth of the primitive reservoir", stressed the evidence from the trace-element ratios Nb/U and Ce/Pb, which are tracers of the formation of continental crust. (Niobium is preferentially retained in the mantle, and lead is transferred to the crust, so both these ratios are higher in mantle that has undergone the extraction of continental crust.) These ratios are identical in mid-ocean-ridge and ocean island basalts, showing not only that both reservoirs have had continental crust extracted from them, but that they must have mixed with each other since most of the continental crust formed.

Harmon Craig (Scripps Institution of Oceanography) was the only speaker willing to stake a claim to having seen primitive mantle. New isotope data for basalts from the Samoan hotspot and the Lau back-arc basin show a correlation between $^3\text{He}/^4\text{He}$, $^{87}\text{Sr}/^{86}\text{Sr}$ and (in the case of Samoa) $^{143}\text{Nd}/^{144}\text{Nd}$. Craig showed that these correlations are consistent with mixing of enriched (Samoa) or depleted (Lau basin) mantle with a primitive mantle component defined by high $^3\text{He}/^4\text{He}$ and 'bulk Earth' strontium and neodymium isotope ratios. Countering this optimism, Stan Hart (Woods Hole Oceanographic Institution) pointed out that no known high- ^3He hotspots are primitive in their lead isotope composition, and specifically that the Samoan basalt array does not point towards 'bulk-Earth' lead.

If mantle plumes do not sample a primitive lower mantle, then where do they come from, and what gives them their characteristic geophysical and geochemical signatures? From convection modelling, we know that buoyant plumes (like atmospheric 'thermals') originate from a heated lower boundary layer — in the Earth's mantle, this could be either the core-mantle boundary or, if the mantle convects in two layers, the boundary between the upper and lower mantle at 670 km. The origin of plumes is thus inextricably linked to the question of whether the mantle convects in one layer or two.

The fate of slabs of oceanic lithosphere

subducted into the mantle gives a handle on this question: if subducted slabs cross the 670-km discontinuity unhindered, whole-mantle convection seems likely; conversely, in a layered mantle, they might be seen to bend towards the horizontal, or thicken, on reaching this depth. Thorne Lay (University of California, Santa Cruz) concluded that seismology as yet gives no unequivocal evidence either way: there are seismic velocity anomalies (reflecting thermal anomalies) along the projections of slabs into the lower mantle, but these could arise from thermally coupled convection across the 670-km boundary. Tomographic images obtained by different groups have shown both continuous slabs down to 1,000 km depth and horizontal deflection at 670 km. Clearly, seismology cannot yet tell us whether the mantle is layered, but Lay held out hope that one day it might.

Is the geochemical evidence for isolated mantle reservoirs not in itself an indication that the mantle is layered? Geoff Davies (Australian National University) argued that a single-layer mantle in which viscosity increases with depth could support chemical heterogeneities with lifetimes of up to 2,000 million years, as required to explain the isotope geochemistry of ocean island basalts (see also ref. 3). In this model, the lowermost mantle has the highest viscosity, and so will be least well mixed; plumes rising from the core-mantle boundary will therefore plausibly sample the oldest and least depleted material, even though there is no strict separation between the upper and lower mantle.

In fact, mantle convection modelling seems at this stage to be permissive rather than prescriptive. For example, a model incorporating a compositional density contrast at the 670-km discontinuity allows, for a small range of values of this parameter, 'leaky' layered convection (Louise Kellogg, University of California, Davis), in which the upper and lower mantle are generally isolated from one another, but plumes rising from the core-mantle boundary can nevertheless penetrate into the upper mantle. The problem with this solution lies in the narrow range of density contrast for which it applies.

Also, the global plume mass flux is sufficient to replace most of the upper mantle over the age of the Earth, so any initial compositional contrast between upper and lower mantle would be homogenized (Norm Sleep, Stanford University).

The ability of a whole-mantle model to preserve chemical heterogeneities for long times in the interior of the convecting cell depends subtly on the details of the flow⁴; doubts of this ability have led to the idea that heterogeneities are accumulated and preserved in boundary layers. Most of the discussion naturally centred on the possible bottom boundary layers, because these would be tapped by plumes, but Don Turcotte (Cornell University) and Claude Allègre (Inst. Physique du Globe, Paris) reminded the meeting that some of the necessary storage time could be spent in the top boundary layer — that is, in the lithosphere.

Although continental lithosphere cannot in general be subducted, Turcotte advocates 'delamination' of the lower continental lithosphere as a way of returning isotopically enriched material to the mantle. Hofmann, arguing that such continental recycling must be volumetrically negligible, focused instead on the most obvious recycling process — the formation and subduction of oceanic lithosphere. In a model of this process (Ulrich Christensen, MPI, Mainz), about a tenth of the mass of subducted ocean crust segregates to the bottom of the system, where it is stored in a low-viscosity boundary layer, and can then be entrained by plumes rising from this layer (Fig. 2). Hofmann argued that different kinds of subducted ocean crust, with or without accompanying terrestrial or deep-sea sediments, could explain the different isotope signatures observed in ocean island basalts.

But this still leaves the question of what constitutes 'the bottom of the system'. Is it the core-mantle boundary (in which case the subducted slabs may account for some of the seismically observed heterogeneity in the D" layer at the base of the mantle), or the boundary between the upper and lower mantle (the "mesosphere boundary layer" of Allègre and Turcotte)? Do plumes rise from the bottom of the mantle, from the bottom of the upper mantle, or from both?

To answer this question by observations of plumes themselves requires some expectation of what the different types should look like. Five years of theoretical and experimental modelling, mostly at the Australian National University (summarized by Ian Campbell), has yielded a comprehensive picture of the behaviour of plumes in the mantle, including the critically important observation that, as plumes rise, their globular 'heads' entrain substantial quantities of the surrounding mantle (Fig. 1).

This means that the size of a plume head reaching the base of the lithosphere depends on its depth of origin. A plume from the core-mantle boundary would have a head

diameter of about 1,000 km by the time it reached the upper mantle⁵; flattening of the head as it approached the base of the lithosphere would double this figure. By contrast, plume heads originating at 670 km are predicted to grow to only about 300 km in their much briefer rise through the upper mantle. The 2,000–2,500 km lateral extent of major continental flood basalt provinces (which in this model are thought to mark the arrival of a plume head at the base of the lithosphere) thus seems to require that at least some plumes originate at the core–mantle boundary.

The wide range in observed buoyancy flux (from 0.3 to at least 8 Mg s⁻¹; ref. 6) suggested to Allègre (among others) that we might be seeing plumes from two boundary layers: the larger ones from the core–mantle boundary, and the smaller ones from 670 km. This would be possible in a 'leaky' layered mantle; alternatively, a single-layer convecting mantle could have a thermal boundary layer at 670 km, arising from the endothermic breakdown of spinel into perovskite plus MgO⁷. But plumes originating in the latter type of thermal boundary layer would have an excess temperature of only about 50 °C (Sleep), probably too small to give an observable surface expression. The observed excess temperatures fall in a rather narrow range, from 160 to 260 °C, and seem not to be correlated with plume size (Jean-Guy Schilling, University of Rhode Island). The wide range in size must therefore have some other explanation, such as the tapping by some plumes of an already depleted boundary layer (Sleep), or simply the diversity that arises naturally in strongly time-dependent convection (Ulrich Hansen, University of Köln).

A final link between plumes and the core–mantle boundary is the fact that the heat flux carried by plumes (about 6–10 per cent of the Earth's total) neatly matches the heat flux out of the core (Davies). This leads to a rather simple picture of the Earth, in which the plate-tectonic cycle cools the mantle (by the return of cold slabs), and mantle plumes cool the core. In this picture, primordial ³He might be stored in the core, but without experimentally determined metal/silicate partition coefficients for helium, it is impossible to say how likely this is.

It now seems that plumes, far from being 'optional extras' to the main pattern of mantle convection, may be our best clues to that pattern. In this regard, Earth-bound geodynamacists may learn something from the tectonics of Venus, where a thinner and less rigid lithosphere does not hide or suppress the fundamental modes of mantle convection (Gerald Schubert, UCLA). Radar

images from the Magellan spacecraft show two types of surface structure that may mark sites of present or past mantle upwelling: circular to oval 'coronae', 200–1,000 km in diameter, with elevated interiors surrounded by concentric ridges; and quasi-circular 'broad rises', 1,000–3,000 km across. It is too early to say whether these correspond to two different scales of mantle convection (Sean Solomon, MIT), and a more definitive

association of surface features with mantle flow patterns will have to await global high-resolution gravity data, such as could be acquired by Magellan on an extended mission. But someday, Venus may well provide the data that will keep mantle convection modellers on the true path. □

Laura Garwin is Physical Sciences Editor of Nature.

PALAEOANTHROPOLOGY

Time for the last Neanderthals

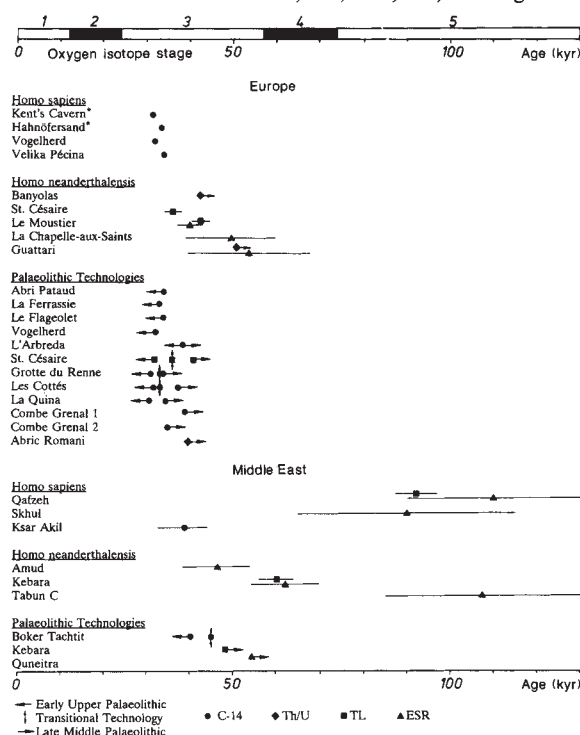
Chris B. Stringer and Rainer Grün

ON page 737 of this issue¹ Mercier and colleagues slot another piece into the jigsaw that constitutes our chronological understanding of the origins of modern *Homo sapiens*. Using thermoluminescence methods, the authors have dated burnt flints associated with a Neanderthal skeleton from Saint-Césaire in France to about 36,000, ± 3,000,

dated Upper Palaeolithic. It was natural to link these two successions, and there was a general assumption that Middle Palaeolithic = Neanderthal and Upper Palaeolithic = early modern. But it gradually emerged that there were actually two early Upper Palaeolithic industries in western Europe — the Aurignacian, which seemed difficult to relate to the preceding local Middle Palaeolithic, and the Châtelperronian, which did indeed have many similarities to the preceding 'Mousterian' of Acheulian Tradition².

Although the Aurignacian could be associated with modern humans at sites such as Cro-Magnon (France) and Vogelherd (Germany), and could reasonably be interpreted as reflecting the immigration of early moderns from further east, the Châtelperronian remained enigmatic. If the Châtelperronian developed from the Mousterian, did its (assumed anatomically modern) manufacturers evolve from Neanderthals? Hints to an answer came from the discovery of Neanderthal-like teeth with the Châtelperronian at Grotte du Renne (Arcy), but it took the discovery at Saint-Césaire in 1979 to finally show that Neanderthals had in fact survived to make the Upper Palaeolithic Châtelperronian^{4,5}. The simplistic equation of hominids and technologies in Europe has thus been abandoned. But since the Châtelperronian and Aurignacian industries can be correlated with each other through stratigraphy and radiocarbon dating, the inferred coexistence of Neanderthals and modern humans has created new explanatory problems. For instance:

- How long was the coexistence phase (or conversely, how rapidly did the Neanderthals disappear)?
- What was the geographical patterning of the coexistence and disappearance: was



Selected age estimates for *Homo sapiens*, *Homo neanderthalensis* and Palaeolithic technologies^{1,7-9,13-20}. *Directly dated hominid specimen^{13,14}.

years before present (BP). This estimate confirms that the fossil is one of the last known Neanderthals, and it supports the proposition that the last Neanderthals coexisted with the first modern humans for several thousand years in Western Europe²⁻⁴.

By the beginning of this century it was apparent that there were two major human successions or transitions during the last glacial stage in Europe: one was from the archaic Neanderthals to early modern humans ('Cro-Magnons'), the other from simpler Middle Palaeolithic (Mousterian) technologies to those of the more sophisti-

1. Wilson, J. T. *Phil. Trans. R. Soc. A* **258**, 145–167 (1965).
2. Morgan, W. J. *Nature* **230**, 42–43 (1971).
3. Davies, G. F. *Earth planet. Sci. Lett.* **99**, 94–109 (1990).
4. Christensen, U. *Earth planet. Sci. Lett.* **95**, 382–394 (1989).
5. Griffiths, R. W. & Campbell, I. H. *Earth planet. Sci. Lett.* **99**, 66–78 (1990).
6. Sleep, N. H. *J. geophys. Res.* **95**, 6715–6736 (1990).
7. Machetel, P. & Weber, P. *Nature* **350**, 55–57 (1991).

there a phased arrival of the Aurignacian in the west, consistent with a migration from elsewhere, and a correlated phased replacement or absorption of the Neanderthals?

- What internal morphological changes can be detected in the last Neanderthal and first modern populations which might reflect evolutionary change or gene flow between them (evidence of hybridization)?

- What behavioural changes can be detected in the late Middle and earliest Upper Palaeolithic archaeological records which might reflect changing adaptations, or interaction between Neanderthal and early modern groups, such as economic competition or acculturation?

Addressing these questions requires above all a better chronological framework for Europe between 30,000 and 50,000 years (30–50 kyr) BP. Part of this time span is beyond the present capabilities of radiocarbon dating, but the newly applied or newly developed dating techniques of uranium series, thermoluminescence (TL) and electron spin resonance (ESR) have the potential to date materials associated with human occupation during this critical period⁶. In the past two years, realization of that potential has grown with age estimates being obtained for five important Neanderthal hominids (Grotta Guattari, La Chapelle-aux-Saints, Le Moustier, Saint-Césaire and Banyolas)^{1,7,8}, as well as for some important archaeological sites. A picture is beginning to emerge of a complex and dynamic period of transition at the interface between the last Neanderthals and first modern humans in Europe.

In the Middle East, things are even more complicated. Although the technological transition from the Middle to the Upper Palaeolithic has been dated to about 45 kyr BP (some 5 kyr earlier than in Western Europe)⁹, TL and ESR dating have supported other lines of evidence that anatomically modern humans were already present about 100 kyr ago^{7,10} (see figure). There seems to be a period during which this area was occupied by both Neanderthals and early modern humans using indistinguishable Middle Palaeolithic technologies. Whether there was a contemporaneous coexistence or a migrational succession of these different hominids remains unknown.

The new technique of uranium-series dating with mass spectrometry has already achieved age estimates with errors of less than 1 per cent in the 100,000 years region when applied to corals¹¹. Applications to secondary carbonates at hominid sites will, one hopes, provide results with similar precision. Thermoluminescence dating of burnt flint already seems to be at a stage where it is providing results with surprisingly small errors. ESR dating of tooth enamel is still fraught with the specific problem of uranium uptake which often prevents precise estimations⁷. Such has been the march of dating technology that we can expect the latter techniques to be further developed and their

reliability enhanced by intercomparison with independent dating methods (such as radiocarbon and uranium series). With that, we should be able to arrive at a more accurate record of the dispersal of modern humans to all parts of the world and the transition from Middle to Upper Palaeolithic technologies.

But what of the Neanderthals? Despite talk of Palaeolithic genocide, it is now clear that their replacement by modern humans was not an overnight process, and the extent and manner of the interactions between Neanderthal and early modern populations could have taken different forms in different areas. Palaeoanthropologists are not agreed as to whether Neanderthals and early modern humans represented different species, but they were probably sufficiently closely related to allow hybridization. Mitochondrial DNA studies suggest that there is no trace of a genetic input from Neanderthal females in recent European samples¹². This, however, is not the same as saying there was no hybridization, although interpreting the patchy fossil record of this period for 'hybrids' is very problematic. The archaeological data from industries such as the Châtelperronian in Western Europe and the Szeletian further east are more suggestive of population interactions^{2,3}, but these are replaced locally by industries firmly associated with modern humans.

The real fate of the Neanderthals may well have been gradual displacement to more marginal and less favourable environments, where their dwindling numbers would have suffered greater attrition from the vagaries of fluctuating climates and food supplies, as well as disease. The Neanderthals probably went out with a whimper, not a bang. □

Chris B. Stringer is in the Human Origins Group, Department of Palaeontology, Natural History Museum, London SW7 5BD, UK. Rainer Grün is in the Subdepartment of Quaternary Research, University of Cambridge, Cambridge CB2 3RS, UK.

1. Mercier, N. et al. *Nature* **351**, 737–739 (1991).
2. Mellars, P. & Stringer, C. B. (eds) *The Human Revolution* (Edinburgh Univ. Press, 1989).
3. Straus, L. G. *Nature* **342**, 476–477 (1989).
4. ApSimon, A. M. *Nature* **287**, 271–272 (1980).
5. Lévêque, F. & Vandermeersch, B. *C. r. hebdom. Seanc. Acad. Sci., Paris*, **291D**, 187–189 (1980).
6. Aitken, M. J. *Science-based Dating in Archaeology* (Longman, London, 1990).
7. Grün, R. & Stringer, C. B. *Archaeometry* (in the press).
8. Julia, R. & Bischoff, J. L. *J. arch. Sci.* (in the press).
9. Marks, A. E. *Préhistoire du Levant* (eds Cauvin, J. & Sanlaville, P.) 287–298 (CNRS, Paris, 1981).
10. Stringer, C. B. *Nature* **331**, 565–566 (1988).
11. Edwards, R. L. et al. *Earth planet. Sci. Lett.* **81**, 175 (1987).
12. Stoneking, M. & Cann, R. *The Human Revolution* (eds Mellars, P. & Stringer, C. B.) 17–30 (Edinburgh Univ. Press, 1989).
13. Hedges, R. E. M. et al. *Archaeometry* **31**, 207 (1989).
14. Berger, R. & Protsch, R. *Radiocarbon* **31**, 55–67 (1989).
15. Gamble, C. *The Palaeolithic Settlement of Europe* (Cambridge Univ. Press, 1986).
16. Smith, F. H. *Curr. Anthropol.* **23**, 667–703 (1982).
17. Mellars, P. A. & Bricker, H. M. *Archaeological Results from Accelerator Dating* (eds Gower, J. A. J. & Hedges, R. E. M.) 73–80 (Oxford Comm. Archaeol. Monogr. 11, 1986).
18. Bischoff, J. L. et al. *J. arch. Sci.* **16**, 563–576 (1989).
19. Bergman, C. A. & Stringer, C. B. *Paléorient* **15**, 99–111 (1989).
20. Ziaei, M. et al. *Quneitra: A Mousterian Site on the Golan Heights* (ed. Goren, N.) 232–235 (Hebrew Univ. Jerusalem, 1990).

Tyre everlasting

THE pneumatic rubber tyre is the crucial invention behind modern transport. Tyres are made in enormous quantities — and later discarded in enormous and intractable heaps. Daedalus now has a self-retreading tyre that will never wear out.

He points out that pneumatic tyres always leak. In the old days, you had to pump up your car's tyres every time you went for petrol. Rubber these days is much more retentive; but it too slowly leaks. Daedalus plans to make use of this effect.

Modern synthetic rubber is largely a polymer of the gas butadiene and the liquid styrene. If a tyre was pumped up with these fluids, they would gradually diffuse through the body of the tyre and escape from its surface. Now the oxygen of the air catalyses the polymerization of such compounds. So on reaching the outside of the tyre they should polymerize, and form new rubber exactly where you want it — on the tyre surface.

Even Daedalus didn't take the idea seriously at first. Oxygen is a very feeble polymerization catalyst. He imagined his novel tyre slowly acquiring a useless sticky surface of partially-polymerized rubber. But he then realized that as the monomers diffused out, oxygen would diffuse in, down an almost equally steep concentration gradient. The polymerization would therefore take place just beneath the surface of the tyre. Even if it took days or weeks, it should be complete long before wear had exposed that layer of rubber.

Even better, this cunning auto-retreading process is self-stabilizing. A rapidly-wearing tyre will get thinner, the monomers will diffuse out faster, and the rate of polymerization will increase. An unevenly worn area (from a serious skid, say) will likewise thicken faster than the rest of the periphery: the tyre will be self-balancing as well. Any actual leak would be rapidly sealed by outflowing monomer, making the new tyre puncture-proof too.

A lot of work will be needed to optimize this complex chemistry. But DREADCO's 'Grotty' should then take over the market. Motorists and environmentalists will rejoice. Even the tyre industry should rapidly adapt to the change. Instead of selling tyres as such, it will sell the monomer fluid to filling stations, to put in their tyre-inflating compressors.

Grotty technology will work on a small scale too. Running shoes with air-cushion inserts could also be pumped up with rubber monomers to counter their ever-thinning soles. But most running shoes are owned by fashion-dazed couch potatoes who seldom even amble. DREADCO's 'Grossoles' will be more useful on the boring, utilitarian shoes of people who really use their feet — housewives, salesmen and Jehovah's witnesses.

David Jones

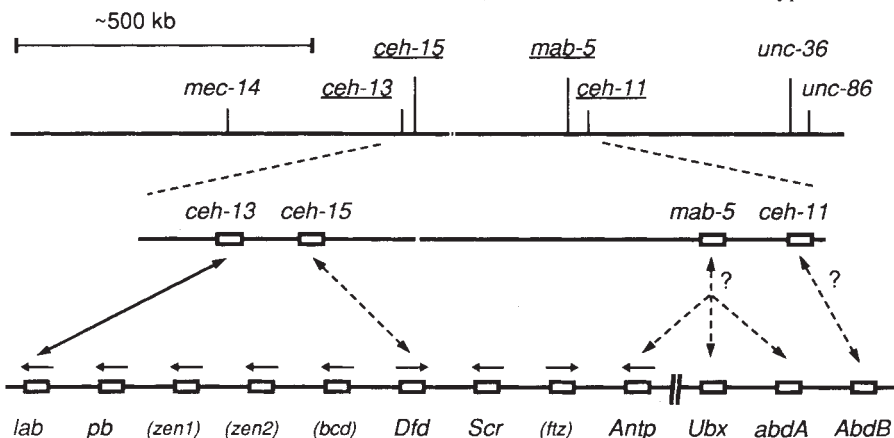
Nematode homeobox cluster

SIR — Clusters of *Antennapedia*-like homeobox genes have been found in *Drosophila* and in vertebrates (see ref. 1 for review) that show a remarkable similarity of organization and expression. Intriguingly, the order of the homeobox genes along the chromosome corresponds to the order of expression domains along the anterior-posterior axis of the animal.

A large-scale screening for homeobox-containing genes² in the nematode *Caenorhabditis elegans* did not initially reveal such clusters. But, as the *C. elegans* map has been

mutants affect cell identities in posterior regions of *C. elegans*, causing transformations to adjacent anterior cell identities.

If the order of homeobox genes along the chromosome were conserved between nematodes, flies and vertebrates, one might expect an *Abd-B*-type homeobox at the location of *ceh-11*. The homeobox of *ceh-11* is highly divergent from all the other *Antp*-like genes (maximum identity of only 57% to *Antp* and 53% to *Abd-B*). Nonetheless, it shares amino acids at certain positions with fly and/or vertebrate *Abd-B*-type homeo-



Physical map of the central region of chromosome III of *C. elegans*. Genes are indicated and the homeobox genes of the cluster are underlined: *ceh-13* and *ceh-15* are about 50 kb from each other, as are *mab-5* and *ceh-11*. *ceh-15* is about 250 kb to the left of *mab-5*; the precise distance cannot yet be determined, as the physical link between these two genes is still tentative in one position where only YAC clones containing repetitive sequences at their ends provide an overlap. The cluster is schematically expanded underneath and compared to the *Drosophila* cluster^{1,9} which is split into two complexes. Direction of transcription for the *Drosophila* homeobox genes is indicated. Genes in parentheses are not involved in anterior-posterior pattern formation. Non-homeobox containing genes in the *Drosophila* cluster have been omitted.

further refined³, a previously unmapped small contiguous set of cosmid clones (locus No. 38 in ref. 2) containing two homeobox genes, *ceh-13* (ref. 4) and *ceh-15* (hom-2/3 in ref. 5; C. Kenyon, personal communication) has now been located on chromosome III, to the left of the homeobox genes *mab-5* (ref. 6) and *ceh-11* (refs 4, 7; see figure). These four genes form a cluster of *Antennapedia*-like homeobox genes that appear to maintain the same anterior-posterior ordering found in flies and vertebrates.

The leftmost gene in the *C. elegans* cluster, *ceh-13*, has a homeobox of the *labial*-type (68% identical to *labial*), while that of *ceh-15* appears most similar to *Deformed*-type homeoboxes (77% identical). The homeobox of *mab-5* is 70–73% identical to a group of closely related (77–93% identity) *Antp*-like genes, that is, *Scr*, *ftz*, *Antp* (the best match), *Ubx* and *abdA*. Flanking sequences like the hexapeptide upstream of the *mab-5* homeobox (VFPWMK) do not aid in assigning *mab-5* to a particular type. It is thus conceivable that *mab-5* is the homologue of some ancestor of several fly homeobox genes. Like members of homeobox clusters in other species, *mab-5* has been shown to control cell identities in particular spatial domains⁶. The *mab-5*

boxes which are not found at those positions in any of the other cluster homeoboxes. In fact, human homeobox clusters each contain five *Abd*-like homeobox genes⁸, two of which are quite divergent. Thus either *ceh-11* is a divergent homologue of *Abd-B*-type homeoboxes, or perhaps a true *Abd-B* homologue remains to be detected. Genetic map position suggests that *ceh-11* could correspond to the gene *egl-5* (refs 4, 7), which functions posteriorly to *mab-5*.

Because *ceh-13* is clearly most similar to the *labial*-type genes that are expressed in anterior regions of fly and vertebrate embryos, and genetic analysis of the more distal *mab-5* show its involvement in pos-

terior pattern formation⁶, we conclude that these genes represent the evolutionary equivalent of the *Antennapedia*-type homeobox gene cluster in higher organisms. Further experiments should reveal whether there are other homeobox genes in this region, and what the most likely evolutionary relationship is between the *C. elegans* homeobox gene cluster and clusters in other organisms. Complete elucidation of the cluster structure will emerge from the *C. elegans* genome sequencing project, since these genes are part of the initial 3-megabase region to be sequenced.

T. R. BURGLIN
G. RUVKUN

Harvard Medical School

A. COULSON

MRC, Cambridge, UK

N. C. HAWKINS
J. D. MCGHEE

University of Calgary

D. SCHALLER
C. WITTMANN
F. MÜLLER

University of Fribourg

R. H. WATERSTON

Washington University, St Louis

(Full addresses can be found in the citations listed.)

Cannabis and night vision

SIR — *Cannabis sativa* is a controlled substance in Jamaica and its unauthorized use is of course illegal. But most fishermen in Kingston and its environs in Jamaica do smoke cannabis or drink an alcoholic extraction of mainly the green stems and leaves of the plant. The alcohol in this case is Jamaica's white rum. The fishermen travel long distances at night in open canoe-type fibre-glass boats. These boats often carry no lighting or compass, and for many years the men have claimed that their vision at night is much better after taking the rum-cannabis extract.

I decided to test this belief, and went with a crew on a dark night to a fishing cay approximately 40 miles south of Kingston. The approaches to the cay are shallow with an abundance of coral reefs, only a narrow entrance of deep water allowing boats to get close for mooring. I sat in the boat and listened for the breakers on the reef but heard nothing, only to be told a short while later that the boat was being docked. At daybreak it was impossible to believe that anyone could navigate a boat without compass and without light in such treacherous surroundings. I was then convinced that the man who had taken the rum-extract of cannabis had far better night vision than I had, and that a subjective effect was not responsible. Note that the fishermen allowed about half-an-hour to an hour to elapse before setting out to sea after taking the extract. And I was told that the effect on

- Gaunt, S. J. & Singh, P. B. *Trends Genet.* **6**, 208–212 (1990).
- Burglin, T. R., Finney, M., Coulson, A. & Ruvkun, G. *Nature* **341**, 239–243 (1989).
- Coulson, A., Waterston, R., Kiff, J., Sulston, J. & Kohara, Y. *Nature* **335**, 184–186 (1988).
- Schaller, D., Wittmann, C., Spicher, A., Müller, F. & Tobler, H. *Nucleic Acids Res.* **18**, 2033–2036 (1990).
- Kamb, A., Weir, M., Rudy, B., Varmus, H. & Kenyon, C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4372–4376 (1989).
- Costa, M., Weir, M., Coulson, A., Sulston, J. & Kenyon, C. *Cell* **55**, 747–756 (1988).
- Hawkins, N. C. & McGhee, J. D. *Nucleic Acids Res.* **18**, 6101–6106 (1990).
- Simeone, A. et al. *Mech. Devel.* **33**, 215–228 (1991).
- Kaufman, T. C., Seeger, M. A. & Olsen, G. *Adv. Genet.* **27**, 309–362 (1990).

their vision was the same whether they smoked the cannabis or took the extract.

I subsequently arranged a meeting with an ophthalmologist colleague, Dr Albert Lockhart, to discuss the possibility of there being a scientific basis to the claimed improvement to night vision. We were able to show that the effect must be due to some component of the cannabis and not the rum. After some years of work, which except in its later stages was carried out with only very limited personal funds and the University of the West Indies laboratory facilities, we prepared a non-psychoactive substance from cannabis, which showed a marked ocular hypotensive effect. This preparation, Canasol, is now used to treat glaucoma, and many patients have reported significant improvement in night vision after taking it.

Lockhart had observed that the incidence of glaucoma seemed to be lower among the rastafarians in Jamaica. The members of this religious group use cannabis in many forms as part of their rituals, and it is known that lowering of intra-ocular pressure is related to cannabis use.

We turned our attention to the possible site and mechanism of this effect. There is an abundance of literature on the pharmacology of *C. sativa*. From our findings, some of which are unpublished, it would seem that adrenoceptors are involved and that they may be located in the ciliary epithelium. The ocular responses can be antagonized by alpha-adrenoceptor blocking agents, ligation of the ascending cervical (sympathetic) nerve, and retrobulbar anaesthesia. To date the mechanism of the improved night vision has not been elucidated. Although lack of money has prevented us from collaborating with ophthalmologists and pharmacologists outside Jamaica, our research continues.

M. E. WEST

Department of Pharmacology,
Faculty of Medical Sciences,
University of the West Indies,
Mona Campus,
Kingston 7,
Jamaica

Too much Hg

SIR — Basu *et al.*¹, in discussing mercury released from crematoria, add 201,780 g plus 120,990 g plus 872,570 g to make 482 kg. But these numbers actually add up to 1195.34 kg. This means that all their following calculations are out by a factor of 2.48.

So, for the crematorium studied by Mills², the mass of mercury released should be $2.199 \times 2.48 = 5.453$ kg per year. This is half-way to the 11 kg he estimated.

V. J. BURTON

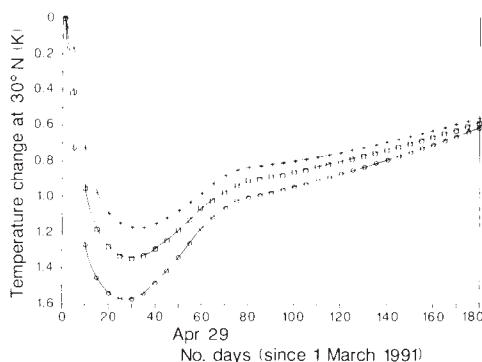
7 St Christopher's Road,
Wildmill,
Bridgend,
Mid-Glamorgan CF31 1RV, UK

1. Basu, M. K., Wilson, H. J. & Krishnan, G. *Nature* **349**, 109 (1991).

2. Mills, A. *Nature* **346**, 615 (1990).

Gulf oil spill and the monsoon

SIR — Attention throughout the world has been focused on one of the biggest environmental disasters caused by the large oil slick and the smoke from the burning of 600 oil wells as a result of the recent Gulf crisis.



Surface cooling over the Indian subcontinent at 30°N during spring and summer 1981. Crosses, Tibet (90°E); squares, India (80°E); circles, Pakistan (70°E).

From a detailed analysis of satellite imageries from INSAT-1B, NOAA-11 and IRS-1A, we conclude that the prolonged smoke from the burning oil wells will have a weakening effect on the Indian monsoon.

Satellite observations (including radiance measurements) show that the deliberate spillage of about 10 million barrels of oil, which started around 24 January, rapidly expanded during the initial 10-day period due to spreading, ocean dynamics and evaporation, to cover an area of more than 6,000 km² from Mina Saud to Abu Ali along the Saudi coast. It stopped expanding further because by then the oil layer had sufficiently thinned to make surface-tension effects dominant¹.

We have estimated the long-term effect of the continuous injection of more than 20,000 tons of smoke from the burning of more than 2.5 million barrels of oil each day on the surface temperature over the Indian subcontinent, assuming a continuous injection rate into the upper troposphere of 3 per cent. The absorption of solar radiation by smoke and its effect on temperature have been estimated for different models of oil-well capping, using the 'nuclear winter' model and taking into account the reduced injection of smoke into the troposphere with them due to the saturation effect.

Calculations for all realistic predictions indicate that the large-scale smoke injection will produce a surface cooling of about 8 K locally, a figure consistent with the observations. More important from the point of view of the Indian monsoon is the resulting long-

term cooling of about 1.0–1.5 K over the Indian subcontinent, particularly over Pakistan, the Tibetan plateau and Delhi, the first two being critical synoptic thermal systems for driving the monsoon (see figure).

Any reduction in the overall solar radiation reaching the Earth will reduce the land-sea heating contrast. Observations of low-latitude volcanic explosions² and analysis of 100 years of surface temperature data³ show that reduction of temperatures by 0.5 K or more over the 'Pakistan heat low' and the 'Tibetan high' regions has a significant impact on the Indian monsoon. The former will weaken one of the two driving terms for the Somali jet, which brings moisture to India through evaporation over both the Indian ocean and the Arabian sea, resulting in the overall weakening of the monsoon⁴. Similarly, the cooling of about 1 K or more over the Tibetan high and the presence of soot, as confirmed by Monex results⁵, will cause a delay in the onset of the monsoon.

Although the monsoon is governed by several global factors, some of which may compensate for the adverse effects of oil smoke, continuous injection of smoke itself will result in regional cooling which will adversely perturb the onset and performance of the southwest monsoon in 1991 and, possibly, in 1992, unless compensated by other factors.

U. R. RAO

Indian Space Research Organisation,
New Bel Road,
Bangalore 560 094,
India

Boring sperm

SIR — By using the title "Sperm whale", Daedalus' research team emphasizes that it thinks only of aggression as a factor in sperm competition¹. But the race is not to the swift. Having reached the neighbourhood of an ovum, sperm must bore through a mass of cumulus cells. When one gets into the ovum, entry of others is impeded; thus boring efficiency may be more important than swimming speed or aggression.

Cumulus is held together by hyaluronic acid; sperm carry hyaluronidase. Working on the assumption that this is not an accident and that boring is helped by hyaluronidase, we tested² hyaluronidase inhibitors such as partly nitrated hyaluronic acid³ and found that they inhibit fertilization in rabbits. Daedalus' team would be well-advised to try hyaluronidase inhibitors as contraceptives, and hyaluronidase activators in the more aggressive aspects of its work.

N. W. PIRIE

42 Leyton Road,
Harpden AL5 2JB, UK

1. Hoult, D. P. A. *Rev. Fluid Mech.* **4**, 341–368 (1972).

2. Handler, P. J. *Geophys. Res.* **91**, 14475–14490 (1986).

3. Jones, P. D., Wigley, T. M. L. & Kelly, P. M. *Mon. Weath. Rev.* **110**, 59–69 (1982).

4. Das, P. K. *The Monsoons* (National Book Trust, New Delhi, 1968).

5. Hamilton M. G. *Proc Int. Conf. Scient. Results Monex 6-21* (1981).

1. Jones, D. *Nature* **351**, 276 (1991).

2. Pincus, G., Pirie, N. W. & Chang, M. C. *Archs Biochem.* **19**, 388–396 (1948).

3. Hadidian, Z. & Pirie, N. W. *Biochem. J.* **42**, 260–265; 266–274 (1948).

Religious resonances

Charles Tanford

Michael Faraday: Sandemanian & Scientist. By Geoffrey Cantor. Macmillan: 1991. Pp.359. £40, \$45.

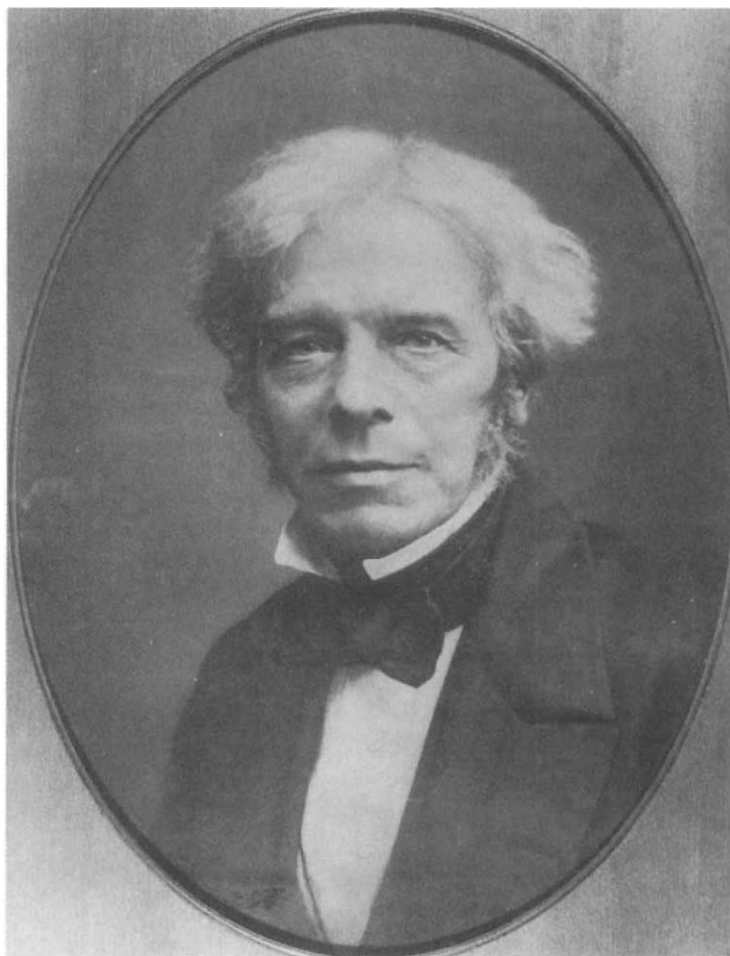
It is rugged country up in the northern Pennines, from where the Faradays came. It is also one of few places in England where the Sandemanian religion gained some measure of a foothold — an appropriately rugged form of Christianity, uncompromisingly guided by scriptural precepts, where a single perceived infraction of the rules led to temporary 'exclusion' and two 'exclusions' meant 'out' forever. And the rules were severe. It was not enough to obey the explicit precepts of the Bible. One had to go further than that, to be guided by these precepts alone, to avoid as far as possible any actions on which the scriptures are silent. Which of course put Sandemanians in opposition to virtually all other religious sects, conformist and non-conformist alike. They had few friends. Michael Faraday himself called it "this despicable sect".

James Faraday was one of the Pennine Sandemanians. He had a smithy in Outhgill, next to an inn on the drovers' trail to market. When an economic downturn stemmed the flow of cattle, he and his wife moved (in early 1791) to London and left the rugged land behind, but took the religion with them. Fellow Sandemanians were in fact instrumental in helping James find a city job. Michael was born in London later that year and he in turn became a fervent Sandemanian for most of his adult life. It is quite appropriate in this Faraday bicentenary year to ponder the question — how much did the demanding constraints of his religion influence Michael Faraday's science? In *Michael Faraday: Sandemanian and Scientist*, Geoffrey Cantor addresses this question — or, to be more precise, he assumes it must have had a direct effect and seeks out evidence to support this.

There are some obstacles. There is no question about Faraday's faith and that it guided him when addressing moral issues, but Faraday himself never claimed a connection with his scientific work and indeed emphasized the need to separate religious and ordinary beliefs in an 1854 lecture. The biblical precepts themselves seem to lead

in this direction — Cantor makes much of Matthew 22:21, the injunction to "render unto Caesar the things which are Caesar's". Does not that advocate a clean separation between spiritual life and one's work in the real world?

Nor does the purely scientific record suggest otherwise. Christian Oersted, who was



Michael Faraday — did his faith influence his scientific work?

the first to see the magnetic needle move in response to an electric current, insisted that he was inspired, not in his case by religion but by the metaphysics of Kant and Schelling. His claim of inspiration can be supported, despite the fact that he seems to have totally misread the kantian philosophy itself. But those who were excited by Oersted's result and followed it up — Ampère, Davy, Faraday — do they not reflect the ordinary progression of discovery, each successive investigator building on the work of his predecessor, driven by the fascination of the problem alone? Or am I being too Whiggish?

Obstacles, however, do not deter Cantor.

He is quite frank about his goal, which is not to make an open-and-shut case but to search for places in Faraday's writings where what he calls "resonances with appropriate biblical passages" can be detected. Likewise he has combed the diaries of Faraday's beloved nieces (Faraday had no children of his own) and the writings of suitable contemporaries, on the basis of whose works Faraday's own thoughts might be "reconstructed".

Whatever the merits of resonances and reconstructions, Cantor's biography does lead to reflection. How about other scientists and other religions? Joseph Priestley, for example: does Unitarianism lurk behind the discovery of oxygen? Or, within an individual scientific life, what are the relative strengths of different social influences? In Faraday's case, for example, the death of the by then rather hostile Humphry Davy, in 1829, as compared with his temporary exclusion from the Sandemanians in 1844 — which had the greater influence?

But we digress, for such comparisons are beyond the scope of Cantor's book. His mission is to gather in one place every conceivable shred of evidence related to his single chosen example — a kind of theoscientific analogue to a clinical case history. He concludes that "no strong causal arrow can be drawn" from Faraday's religion to his science or vice versa. Both were "very much of a piece", he states, both "chosen responses to his psychological needs".

My own thoughts turned to my former chemistry mentor, Henry Eyring. He was a devout Mormon, an elder of the church's governing body in Salt Lake City. Scriptural precepts dominated his every moral stance and his family life. But none of us who were his students could imagine any relation to his science

(which included a brilliant book, *The Theory of Rate Processes*; McGraw-Hill, 1941). Were we naive? I must find myself a copy of the Book of Mormon and search for resonances! □

Charles Tanford is at Tarlwood, Back Lane, Easingwold, North Yorkshire, YO6 3BG, UK.

■ Also published recently to commemorate the bicentenary of Faraday's birth is *The Correspondence of Michael Faraday: Vol. 1, 1811-1831*, edited by Frank A. J. L. James, which contains many previously unpublished letters both to and from great British scientists. Published by The Institution of Electrical Engineers, price is £55. □

Grit and grain

Robert S. Anderson

Aeolian Sand and Sand Dunes. By Kenneth Pye and Haim Tsoar. *Hyman*: 1990. Pp.396. £70.

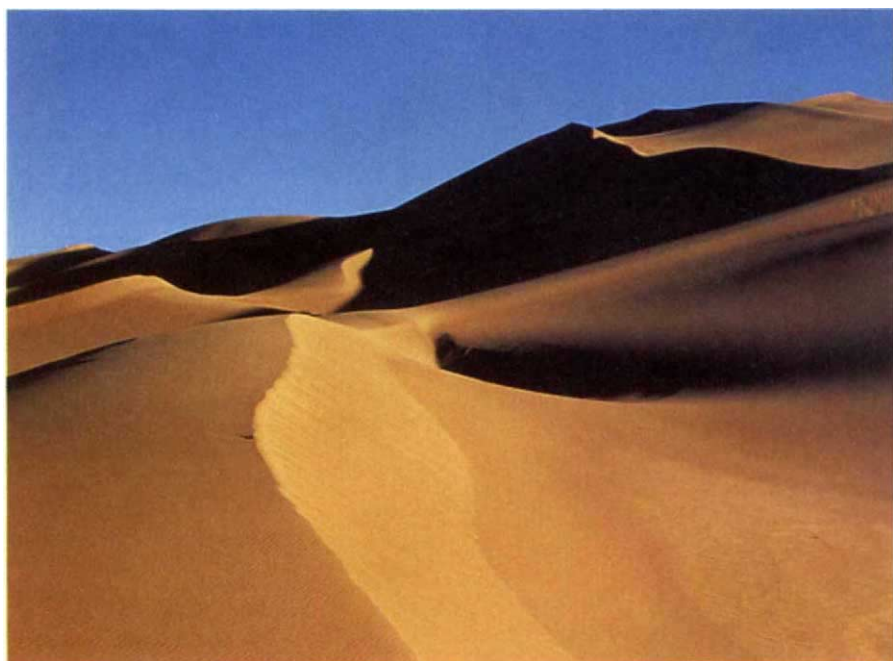
THE world has recently witnessed a war fought along and above the sandy desert surfaces of Iraq and Kuwait, a war waged, ultimately, for control over vast oil reserves trapped at depth in porous sandstone and carbonate reservoirs. Throughout the world, people produce crops grown in a granular substrate that has a nasty habit of being blown away by the wind, and whose success or failure depends upon the vicissitudes of climate. Although deep-sea deposits com-

transported, the mechanics of the transport process, the deposits they engender, the postdepositional modification of the sediments, and the history of climate one may infer from a study of the deposits.

This is the second of two recent attempts to collect together the aeolian literature. The first, R. Greeley and J. Iversen's *Wind as a Geological Process* (Cambridge University Press, 1985), sought to review the literature to provide a context in which to interpret wind transport systems on other planets. Pye and Tsoar emphasize instead the sedimentological techniques and problems, and on these fronts they largely succeed. We learn about the history of the coastal dunefield of Oregon, about the Simpson–Strzelecki sand sea of central Australia, and about the post-depositional cementation processes. Also discussed is the evidence used to infer that

also parochial in the sense that they have not incorporated extensive developments in our understanding of aqueous sediment transport over the past 20 years that are indeed pertinent to an understanding of aeolian sediment transport and associated bedforms. In their defence, however, this is also one of the faster-growing portions of the literature. The physics community has recently become engaged by treating both the saltation transport up the windward sides of dunes and the grain flows down the lee sides of sand dunes as examples of a general class of problems in granular mechanics. The most recent work will appear soon in two special issues of *Acta Mechanica*. □

Robert S. Anderson is in the Department of Earth Sciences, University of California, Santa Cruz, California 95064, USA.



Science Photo Library

Shifting sands — human interaction with a sandy substrate goes back to the Middle Ages.

prise the principal source of our climate record, the distribution of aeolian (windblown) deposits provides an important component of the terrestrial record, from which one wishes to infer the histories of wind speeds, directions and vegetational resistance to sand transport. It is therefore of increasing interest to understand the processes involved in the transport, deposition and lithification of sand, and to amass case histories of climate from such deposits.

Kenneth Pye and Haim Tsoar, both with extensive experience in aeolian matters, have teamed up to produce *Aeolian Sand and Sand Dunes*, a book whose primary purpose is to collect the disparate and growing literature on the topic. The book may be viewed as a sequel to Pye's first book, *Aeolian Dust and Dust Deposits* (Academic, 1987). Both serve as large-scale literature reviews, introducing the reader to recent work on the nature of the materials, the conditions under which these grains will be

the extent of the sand seas at the height of the latest glaciation was quite different from that at present. And there is a sufficient review of atmospheric circulation to allow an understanding of how this might have come about. The authors include much information covering the practical details of field and laboratory measurements used to characterize both the sand and its transport. In concluding, Pye and Tsoar address the larger issues of human interaction with a mobile, sandy substrate, an interaction with a rich history reaching back into the Middle Ages.

The weakest part of the book is its treatment of the mechanisms of sediment transport. Neither of the authors are mechanistic, and their review of this literature is neither complete nor as critical as it should be. Although much has been added to our understanding of aeolian sediment transport mechanics since Greeley and Iversen's book, this is not well reviewed. Their treatment is

The bangs of a magician's brain

Alvaro De Rujula

The Charm of Physics. By Sheldon L. Glashow. *American Institute of Physics*: 1991. Pp.300. \$27.95, £17.50.

GREAT steps in our knowledge of physical reality often occur in the form of a 'unification': an understanding of different concepts or phenomena as aspects of the same underlying principles. No doubt the greatest unification of all times is that due to Newton: the discovery that the motions of the Moon, the planets and the fake proverbial apple all obey the same universal law. Competing for second place are Maxwell, for the unification of electricity, magnetism and light, and Einstein, for unifying light (or gravity), space and time. No doubt the Swedish Academy would vote for Maxwell, as Einstein was never awarded the Nobel prize for either of his 'relativities'.

Sheldon L. Glashow, Abdus Salam and Steve Weinberg received the 1979 Nobel prize for the most recent unification, that of electromagnetism and the weak interactions. Perhaps in an attempt to compensate for its earlier disregard of unifications, the Swedish Academy awarded this particular prize before experiment had confirmed the main prediction of the new theory by proving the existence of the W and Z particles that are the 'carriers' of the weak force. The daring Swedes must have held their breath rather tightly for the few years that elapsed until the actual discovery of these particles.

The Charm of Physics is a collection of essays by Sheldon Glashow, many of them appearing in print for the first time. The essays date from 1975 to the present, and have not been edited to exploit the benefit of hindsight. Sometimes they relate to a very particular occasion, and the book is best read with this in mind. For example, the first

chapter, which is based on a speech, gives such meagre information about the many people named that I can only conclude that either they were present at the speech or were known to those who were.

Only a couple of the 30-odd essays deal specifically with the unification of the weak and electromagnetic forces. And only a handful discuss those of the author's predictions that turned out to be correct, such as the existence of the 'charmed quarks' that give the book its name. There seems to be only one set of laws of nature, and there are many fascinating possibilities that nature appears to have ignored. As most creative scientists do, Glashow has stumbled upon many more of the latter than of the former. But unlike many of his colleagues, he is ready to discuss them both. An example of an idea whose experimental verification nature seems mysteriously to be postponing is the 'grand unification' of (almost) all known forces that, as Glashow admits, has so far turned out to be like the Holy Roman Empire, that is, neither grand nor unified.

The Charm of Physics contains autobiographical material, occasional attempts at

poetry, blood-chilling views on nuclear holocaust and other self-imposed threats to mankind, fascinating recollections of the discovery and successive theories of chemical elements, planets, galaxies, pulsars and quarks. Glashow also treats us to delightful introductions to various aspects of mathematics, chemistry, physics and cosmology, including instructions on how to find our distant neighbours, the 'little green people' that, the author elaborately argues, live in large numbers elsewhere in our galaxy. The style is light and humorous, except in those essays that deal with more serious issues, such as the likely fate of mankind or, perhaps more immediate, the future of science and technology — and of education in general — in countries such as Britain or the United States.

Glashow's book, and he himself, are both deep and superficial, serious and hilarious, thoughtful and scatter-brained and much more besides. Although they defy summary, they truly deserve attention. □

Alvaro De Rújula is in the Theory Division, CERN, CH-Genève 23, Switzerland.

Legal thinking

Andrew Grubb

Family Rights: Family Law and Medical Advances. Edited by Elaine Sutherland and Alexander McCall Smith. *Edinburgh University Press: 1991. Pp.135. £25. (Distributed in the US by Columbia University Press, \$39.)*

DURING the past decade there has been a remarkable growth of interest in medical law. One of the principal areas of study and research for medical lawyers has been the legal regulation of decision-making about medical treatment and the family: decision-making concerning when not to have a child (contraception and sterilization); when to create a child (*in vitro*-fertilization treatment and surrogacy); when to end a pregnancy (abortion and selective reduction); what controls to impose on a pregnant woman out of a concern for her unborn child (paternal injunctions prohibiting abortion, court-ordered caesarian sections) and, finally, what treatment an ill child should receive (withholding life-saving treatment from, for example, handicapped babies).

There are few, if any, easy solutions to many of these areas of public controversy. Two principal questions arise: first, should the law regulate the particular area of decision-making and second, if so, what should that regulation be? *Family Rights* is a volume of seven essays, predominantly written by academic lawyers, which addresses many of the legal dilemmas that modern medicine has created for society.

As the essays illustrate, the law is a potent regulatory force in policing the moral dilem-

mas of family decision-making, and rightly so, given the important human interests that are in play. But the law is often slow at providing a response to a new development. Judges can only deal with cases that are brought before them and Parliament may not want to intervene where a morally 'hot potato' exists. As Bernard Dickens shows, until recently the English common law provided almost no mechanism for the regulation of modern reproductive technology except, perhaps, in relation to status questions of parenthood based upon what some would call the approach of a by-gone era. Fortunately, after six years of vacillation, Parliament passed the Human Fertilization and Embryology Act 1990, which provides a comprehensive regulatory framework for modern reproductive technologies. But this is not always the case. Surrogacy, for example, remains unregulated in England despite the 1990 Act. Could it be, as Sheila McLean suggests in her essay, that the moral controversy has not been explored sufficiently for politicians to make up their minds? Certainly, as she points out, few can ignore the commercial element which is often present and which produces such strong feelings of revulsion.

Sometimes judges defer to the legislature because of the controversial nature of any law they would develop. Hence, as Elaine Sutherland catalogues, English judges have been reluctant to become involved in regulating decisions concerning pregnant women. Courts have refused injunctions to fathers who are seeking to prevent an abortion and have also refused to exercise their protective jurisdiction over an unborn child when an attempt is made to protect it from its mother's (alleged) dangerous conduct. The theoretical difficulties of intervening — dis-

regard of the woman's right of autonomy — and the practical problems of enforcement conspire against judicial intervention. Sutherland argues in favour of some legal intervention. Not everyone would agree with this approach, although there is a growing body of persuasive authority in the United States which would support it.

By contrast, in a number of areas judicial intervention has been accepted often out of a perceived need to protect an individual; sterilization of mentally handicapped girls and women, for example. Similarly, judges are willing to override parental decisions in relation to nontreatment of their children, in particular handicapped babies, if they consider the parents' decision not to be in the child's best interests. Douglas Cuisine and David Meyers examine the judicial approaches to these two areas of family decision-making.

Faced with this largely interventionist judicial attitude, it is left to Sandy McCall Smith to challenge its basis and to sound a note of caution. Should the courts 'second-guess' parental decisions or should they leave them to those whom it is usually accepted best understand the needs of their children? This is a fundamental question. It features, rightly, in the first chapter. But on the basis of the English case law, McCall Smith has really lost his argument: courts do, on the whole, intervene as protectors of a child's best interests.

Family Rights will provide the general reader with a reasonably accessible account of current legal thinking. There is undoubtedly more to come from the courts (and Parliament) in the future. The book deserves to be widely read and will set the scene for the public debate, which has really only just begun. □

Andrew Grubb is at the Centre of Medical Law and Ethics, King's College London, Strand, London WC2R 2LS, UK.

New in paperback

■ The paperback edition of *Hormones: From Molecules to Disease* edited by Etienne-Emile Baulieu and Paul A. Kelly (for review see *Nature* **351**, 362; 1991) is available from Chapman and Hall at a price of £47.50, \$62.50. □

■ For an introduction to polymer science suitable for undergraduate courses, see the second edition of *Polymers: Chemistry & Physics of Modern Materials* by J. M. G. Cowie. Published by Blackie/Chapman and Hall, price is £19.95. □

■ *Sequence Analysis Primer* edited by Michael Grobsov and John Devereux provides a useful introduction to computerized sequence analysis in biotechnological research. The book is published by Stockton Press at \$39.95. □

■ Those interested in body size, life histories and evolutionary biology in general should look out for *The Allometry of Growth and Reproduction* by M. J. Reiss. The book is published by Cambridge University Press at a price of £12.50, \$19.95. □

Mixed material

John A. Campbell

Biocomputers: The Next Generation from Japan. Edited by Tsuguchika Kaminuma and Gen Matsumoto. *Chapman & Hall: 1991. Pp.206. £25.95, \$49.95.*

THE first problem for anyone interested in the idea of biocomputing is to discover exactly what biocomputing is. The evidence is likely to be confusing. Neural networks, biosensors, molecular computing and loose descriptions of Japanese 'post-fifth-generation' programmes are only a sample of the ingredients in the mixture. Although it is easy to track down articles about each of them, it has been much more difficult to find material that gives an overall perspective.

This book fills this gap very creditably. Each of the relevant subjects gets proper attention, confirming that biocomputing is indeed still very much of a mixture of topics and ideas waiting for a helping hand towards coherence. Some degree of coherence will certainly be needed if devices that can reasonably be called biocomputers are to emerge from the present variety of theoretical and experimental work. The prospects are promising for the medium term, but commercial payoffs will not happen tomorrow, or even the day after tomorrow. The book's contents show why this is so, although the title oversells the prospects. The usual explanation for such mismatches in books on advanced technologies is overeagerness on the part of a publisher.

At present, biocomputing is a convenient name to cover activities in all the areas that involve research on chemical or biological phenomena, or idealized models of the phenomena, with a bias towards applications in computing. One reason for the diversity of areas is that there is no obvious single or most rewarding target: computing can mean many different things, from limited specialized processing of data by molecular or biological sensors to large-scale imitations of intelligent behaviour in drawing analogies and making complex inferences. A common requirement for progress in all areas, however, is that each type of participant should understand the general shape and the essential results of other participants' fields. In particular, computing specialists need to appreciate the achievements and problems of the participating biological scientists, and vice versa. The book does a good job of setting up an educational agenda and signposts for further reading for computer scientists, although it is rather less effective in explaining the results and preoccupations of advanced computer science to others. Even so, it is the best all-round educational source for biocomputing that exists at the moment.

Some of its information is helpful in making sense of the news of Japanese

research and development programmes that are announced from time to time. For example, the 'sixth generation' label is associated with planned work with organic molecules rather than silicon-based technology. A further Japanese initiative, identified in the book as being about "technological creation of a system which processes information in ways resembling human intelligence", has the dubious distinction of receiving the 'seventh generation' label from Western journalists. These programmes are described accurately, along with several others, in a manner that is more sober than in much Japanese writing on related subjects. The authors appreciate and state the difficulties matter-of-factly; there are only small traces of the purple prose (not unusual in the early days of the 'fifth generation') that fits some of the demands of higher Japanese literary style and is calculated to inspire the scientific troops. The general picture is of a set of overlapping but diffuse scientific objectives — exactly following the actual situation for biocomputing itself, in fact.

Because the subject is even more ripe than 'fifth generation' computing for precompetitive research — a favourite phrase at the beginning of the rival European Strategic Programme for Research and Development in Information Technology — it needs more direct collaboration between industry and universities, and between representatives of

different disciplines, than has occurred in Japan in the past. The editors raise this issue a little ambiguously: they mention collaborations that are already being encouraged, yet also argue for collaboration with almost missionary fervour. To encourage it in some Japanese university and commercial environments requires the combined skills of a missionary and a *samurai*, so the fervour is understandable. The European record on this front is better, but not so much better that we can be pleased about it, especially given the present tendency towards pressing for shorter-term results from programmes in advanced technologies in Europe and decreasing the amount of university participation.

After the serious questions are disposed of, *Biocomputers* is also good for some innocent enjoyment. Apart from a few startling but harmless mistakes in the history of science, the main source of novelties (such as the misappropriation of K. Blodgett's credit with I. Langmuir for their molecular-film deposition method by a mystery physicist called Brogette) is Japanese phonetics. Fortunately, the main messages of the book are trustworthy; disinformation is confined to the decorations around the edges. □

John A. Campbell is in the Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK.

Distant vision

P. G. Whitehead

Remote Sensing in Hydrology. By E. T. Engelman and R. J. Gurney. *Chapman & Hall: 1990. Pp.225. £35.*

REMOTE sensing is often seen as being primarily a series of techniques looking for a problem to solve. The high-tech nature of the business, the endless acronyms and unintelligible technical jargon have created a sub-world (or superworld) of specialists who can speak eloquently to each other but who have trouble communicating with outsiders (that is, those scientists who might actually want to use the technique).

It was a great pleasure, therefore, to read this book, which assumes little knowledge of remote sensing, explains the technology in a clear and concise manner and demonstrates a wide range of applications in one field, that of hydrology. Both authors are experienced hydrologists who have adapted remote sensing to meet their particular needs. The book is perfectly readable to the nonhydrologist, who is briefed in the first chapter on the basics of the hydrological cycle and then later in chapters on precipitation, snow, evapotranspiration, runoff, soil moisture, groundwater and water quality. The distributed nature of hydrological problems such as flooding patterns across regions or estimating snow cover over an inaccessible moun-

tainous area ensures that remote sensing can play an important role. But real operational applications have so far been restricted by practical necessities such as the need for clear weather to ensure a cloud-free image. New technologies based on active radar which can 'see' through cloud systems will enhance the general applicability of remote sensing.

As the authors point out, remote sensing is not yet an operational tool. But the necessary research has been undertaken in many fields within hydrology to demonstrate the potential for operational management. To some extent, this is already happening in the area of flood forecasting, where spatial information on rainfall obtained by ground-based weather radar systems is being routinely used to update flood forecasting models.

Remote sensing is at last offering a reliable series of techniques that can provide hydrologists with new insights, new datasets, regional and global information. There is now a major challenge for hydrologists to use this information to assist in answering questions on the global water balance, soil moisture, vegetation processes, calibration of models and to match the level of large-scale or macro-modelling achieved by atmospheric and ocean modellers. This is perhaps where remote sensing has a key role to play in future years. □

P. G. Whitehead is at the Institute of Hydrology, Maclean Building, Crowmarsh, Gifford, Wallingford, Oxfordshire OX10 8BB, UK.

Visual detection of 1981S13, Saturn's eighteenth satellite, and its role in the Encke gap

Mark R. Showalter

Center for Radar Astronomy, Stanford University, Stanford, California 94305, USA

Careful inspection of many images taken by the Voyager spacecraft reveals the presence within the Encke gap of Saturn's eighteenth satellite. Its existence had been inferred from gravitational disturbances seen in Voyager data, and it falls close to the predicted orbit. Its shepherding effect is responsible for keeping the Encke gap open, and it may also be the progenitor of a narrow ringlet within the gap.

THE Encke gap is the second most prominent division in Saturn's ring system, after the broad Cassini division separating rings A and B. It is 320 km wide and falls in the outermost third of the A ring. Cuzzi and Scargle¹ were the first to notice that the gap edges show a distinctly 'scalloped' appearance in many of the highest-resolution images. The pattern has a wavelength of 0.7° in longitude (1,500 km) and a radial amplitude of a few kilometres. They suggested that the pattern could be explained by the gravitational perturbations of a small moonlet orbiting near the centre of the gap. The body's orbital position could be pinned down to a $\sim 30^\circ$ zone of longitude that was not imaged at high resolution by the Voyager cameras. Showalter *et al.*² followed up this work with a kinematic model for the effects of a moonlet on ring material beyond the gap edges. They demonstrated that the same perturbations will give rise to a quasiperiodic pattern of optical depth variations which remains fixed in a frame orbiting with the moonlet, and they dubbed this pattern a moonlet 'wake' by analogy with the water waves trailing behind a motorboat. These investigators discovered wake-like patterns near the Encke gap in several Voyager data sets, including the Voyager 2 photopolarimeter (PPS) stellar occultation profile³ (Fig. 1) and the Voyager 1 radio science subsystem (RSS) Earth occultation profile⁴. By fitting their model to the observed wakes, Showalter *et al.* derived an accurate orbit for the as yet unseen moonlet. They also inferred a mass which was compatible with the original edge wave analysis¹, corresponding to a moonlet radius of ~ 10 km if it is composed of water ice.

The orbit deduced for the Encke moonlet was subject to a number of uncertainties. The kinematic model of Showalter *et al.*² neglects any collective effects among the ring particles, such as collisions and ring self-gravity, which would slowly alter a wake pattern. By good fortune, this simplification was valid within the best-resolved wake, just inward from the Encke gap in the PPS occultation profile (Fig. 1). The moonlet was calculated to lie a mere 32° away in orbital longitude at this observation, leaving insufficient time for collective effects to have built up to a significant level. Based on the wake, the best-fit semimajor axis a for the orbit is $133,595 \pm 5$ km (after including a recent 7.4-km correction to the radial scale of the PPS profile⁵); this corresponds to an orbital mean motion $\Omega = 625.95^\circ$ per day. This orbital solution was roughly compatible with that derived from the other wake patterns observed in the PPS and RSS data, although the other profiles fall much farther in longitude from

the moonlet and show evidence for modification by collective effects.

Image search procedure

Shortly after the analysis of the wave edge and wake provided convincing evidence for a moonlet in the Encke gap, a cursory search through the images by eye revealed that none of the highest-resolution Voyager images had captured it. But precise knowledge of the body's orbit² makes possible an automated computer search through the entire catalogue of $\sim 30,000$ Voyager images from Saturn. Supplementary data files available to Voyager imaging scientists describe the viewing geometry of each Voyager frame, including the spacecraft position, camera pointing direction and orientation, as well as the time of the exposure. I developed a program to read through this file and calculate, for each Voyager frame, whether the orbiting moonlet was expected to fall within the camera's field of view. As the catalogued pointing information is often imprecise, the search program did not exclude an image unless the moonlet's predicted location was >100 pixels outside the nominal field of view (which is 800×800 pixels). The program performed additional tests to ensure that the spatial resolution was finer than a threshold of 100 km per pixel, and that the body was not obscured or shadowed by Saturn.

As the moonlet's position was most accurately determined for the Voyager 2 epoch, this image catalogue was considered first. The list of images returned by the program included some with promising properties, including low solar phase angles and spatial resolution approaching 30 km per pixel. Figure 2 shows the highest-resolution candidate, 43918.04. A bright spot is indeed visible in the Encke gap, close to the moonlet's predicted

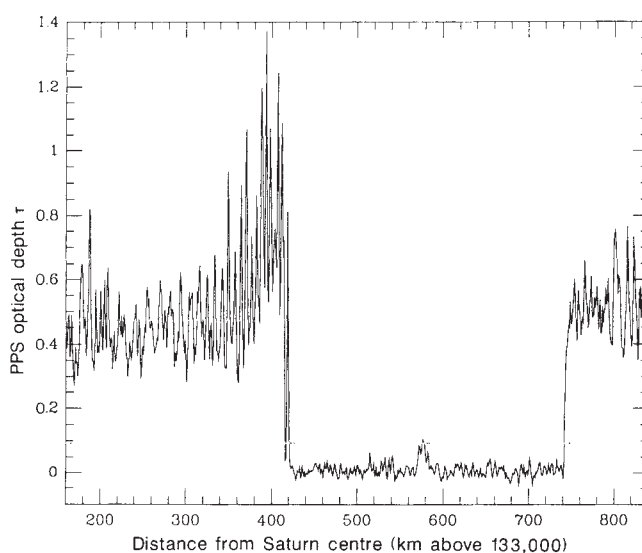


FIG. 1 Optical depth profile across the Encke gap, as obtained from the Voyager 2 PPS scan. The moonlet wake is visible as a quasiperiodic optical depth variation inward (left) of the gap. A faint, narrow ringlet is also visible near the gap's centre.

TABLE 1 Summary of images of 1981S13

Image	Measured location	Resolution (km per pixel)	Σ DN*
34740.42†	(786, 529)	84.4	(8)
43708.39	(713, 614)	91.7	(11)
43752.22	(283, 120)	78.7	(15)
43752.26	(342, 199)	78.6	(14)
43785.17	(383, 411)	69.8	(18)
43785.21	(192, 208)	69.7	(18)
43796.05	(539, 516)	66.8	(41)
43796.09	(404, 413)	66.8	(21)
43796.13	(100, 249)	66.8	(21)
43810.52	(661, 164)	61.4	(24)
43810.56	(401, 6)	61.4	(24)
43818.37	(437, 745)	60.5	(24)
43860.37	(421, 236)	46.1	(42)
43861.44	(110, 300)	46.1	32
43864.56	(767, 768)	46.4	80
43865.00	(426, 565)	46.4	33
43865.04	(89, 321)	46.4	31
43904.59	(213, 745)	34.8	72
43905.05	(188, 693)	34.7	80
43905.11	(135, 605)	34.6	80
43905.17	(77, 529)	34.6	102
43905.23	(25, 431)	34.5	88
43918.00	(122, 411)	30.9	103
43918.04	(725, 537)	30.9	83

* DN stands for data number. Parentheses indicate predicted values for images in which a measurement was not performed.

† Uncertain, marginal Voyager 1 detection.

location. Figure 3 contains a sequence of five images, 43904.59–43905.23, taken a few hours earlier, which show the body crossing the camera's relatively fixed field of view. These images demonstrate that the Encke moonlet was indeed recorded by the Voyager cameras, and they provide a striking confirmation of the wake theory. The body has now been designated 1981S13.

Photometry

In the best images, it is relatively straightforward to measure the disk-integrated brightness of 1981S13. The most direct measurement is simply the sum of DNs across the body, where DN is the integer 'data number' proportional to the brightness in each pixel of an unprocessed Voyager image. When measuring

Σ DN, care was taken to exclude any light from the nearby rings and to subtract out the (typically nonzero) background level. Knowledge of the image's filter properties, exposure time and spatial resolution then permits a direct conversion to a disk-integrated reflectivity. Using measurements from the 10 images with finest spatial resolution (see Table 1), the disk-integrated ratio I/F for 1981S13 is $141 \pm 7 \text{ km}^2$ at a mean phase angle of 10° ; here I is the reflected intensity and πF is the incident solar flux density. Allowing for the uncertainties in its mass and density, the body has a geometric albedo of 0.4–0.6. This albedo range is similar to that of other inner moons of Saturn, such as Pandora, Prometheus and Atlas⁶.

Note that this result provides fairly convincing evidence that 1981S13 has an icy composition. If it consisted primarily of rocky or carbonaceous material, it would be likely to be both darker and smaller (because its density would be higher but its mass unchanged). To fit the observed brightness, 1981S13 must have both a large area and a bright surface; water ice is the only common material that satisfies both of these requirements.

Orbit determination

The visual detection of the Encke moonlet makes it possible to determine its orbit directly, instead of inferring it from a kinematic model for moonlet wakes. I have measured the orbital longitude of 1981S13 in each Voyager image in which it appears. Precise measurements require a pointing correction to be determined empirically for each Voyager frame, to align all visible image features with their predicted locations. I determined the pointing offsets using the observed location of the Encke gap, and other image features as available, including the location of the shadow boundary crossing the rings, Saturn's limb and terminator (the edge of Saturn's disk, and the boundary between day and night on Saturn's surface), and the outer edge of the B ring using a two-lobed model which precesses under the influence of Mimas⁷. In general, the corrected pointing is accurate to a fraction of a pixel. The longitude of the moon is then measured, corrected for light travel time, and precessed to the time of the Voyager 2 PPS scan (assuming the wake-derived mean motion of 625.95° per day). To estimate the uncertainty in each measurement, I have determined the range of longitudes compatible with a one-pixel offset in the position of the moon.

The exhaustive list of images showing 1981S13 (Table 1) was acquired using an iterative procedure, in which a new orbit was fitted to the set of measured longitudes, and the search procedure was re-run to identify other images showing the body. Surpris-

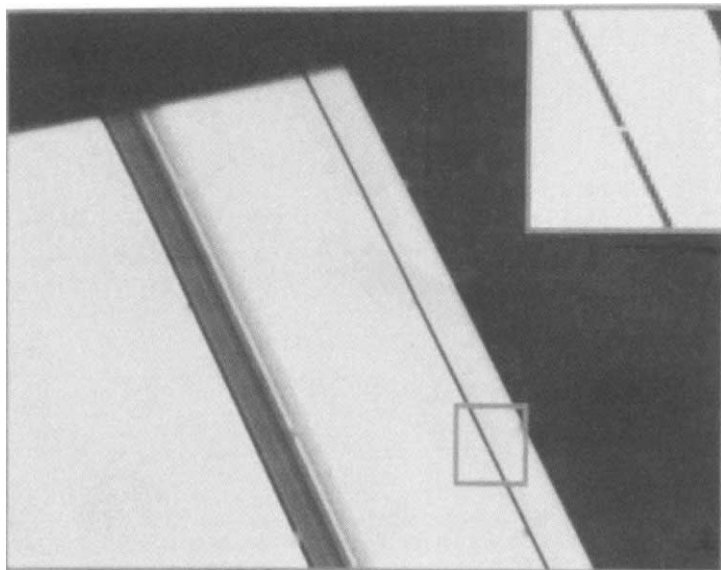


FIG. 2 Voyager 2 image 43918.04 is the highest-resolution image available showing 1981S13. The A ring enters the shadow of Saturn to the left, and a small box identifies the location of the moonlet. The moonlet is shown enlarged and with enhanced contrast in the inset box.

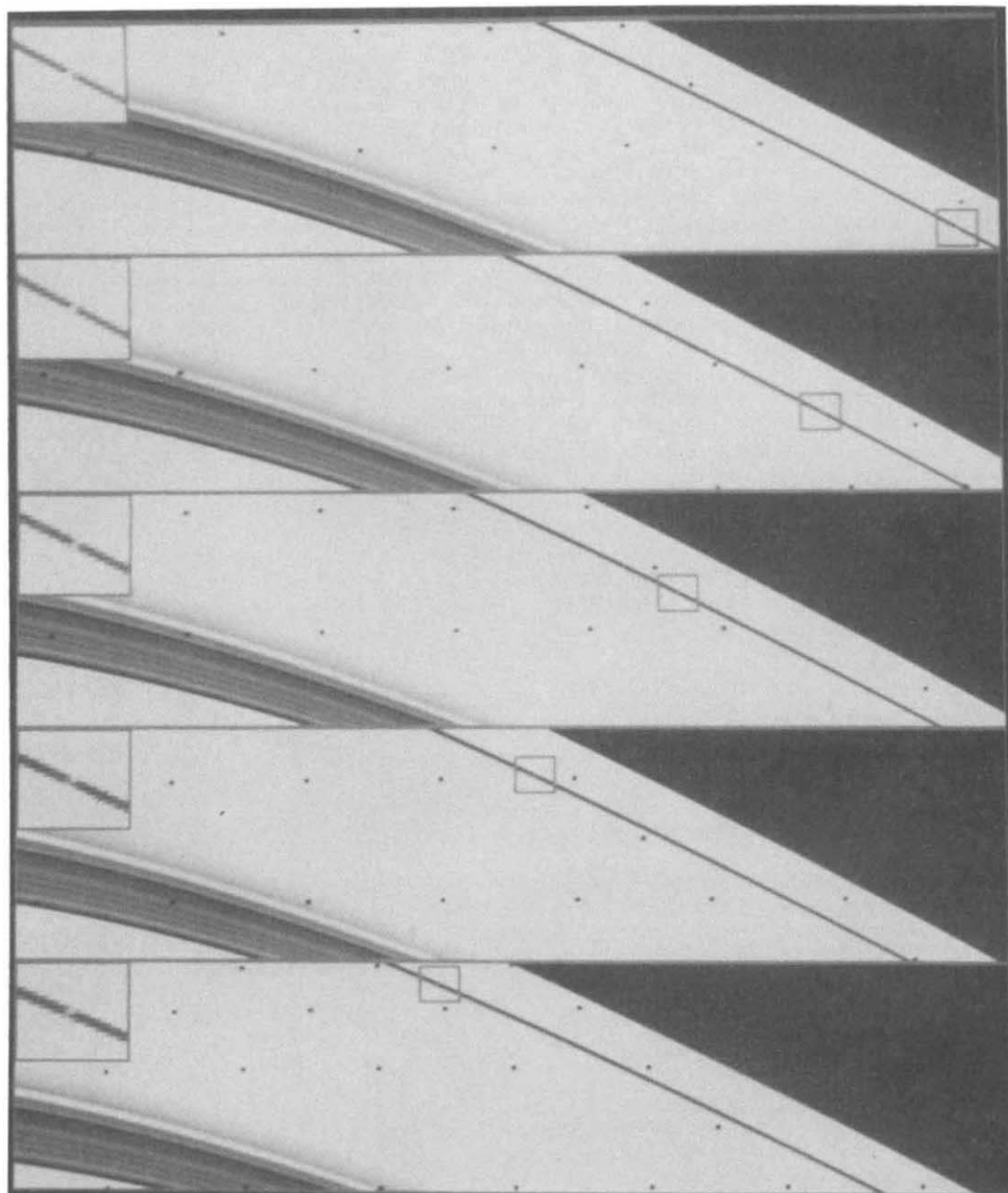
ingly, the moon could be detected in 23 images, with spatial resolution as poor as 91 km per pixel. The observations span an eight-day period before the closest approach, encompassing about 15 full orbits. In general, the body was detected in each Voyager 2 image for which the expected sum of data numbers across the body was ≥ 10 and the projected width of the gap at the moonlet's location was ≥ 3 pixels.

Figure 4 shows the offset in longitude of 1981S13 from its predicted location in each Voyager frame, as a function of the observation time. The nonzero mean value and slope indicate a significant error in the predicted orbit, which will be explained below. The figure also illustrates that the procedure for estimating errors was too conservative, and typical uncertainties in the moonlet's measured longitude are actually closer to 0.5 pixels. The mean motion determined from these direct observations is $\Omega = 626.044 \pm 0.006^\circ$ per day, which corresponds to $a = 133,582.8 \pm 0.8$ km using the most recent models for Saturn's gravity field^{7,8}. At the time of image 43905.11 (19:03:09.2 Universal Time on 22 August 1981), the body lies at ring plane longitude 123.617° , measured relative to the ascending node of

Saturn's ring plane on Earth's mean equator of 1950. The small formal error in the mean motion propagates ahead to a longitude uncertainty of $\pm 45^\circ$ at the time of the Cassini Orbiter's arrival at Saturn after the turn of the century.

Obviously, a single detection by Voyager 1, some 9 months earlier, would greatly improve the accuracy of the inferred mean motion. In applying the search procedure to the Voyager 1 image catalogue orbits up to ± 3 standard deviations from the best fit were considered, to ensure that no images were missed. Nevertheless, the search did not yield promising results. The Voyager 1 sequence contained far fewer views of Saturn's rings, particularly during the approach phase when 1981S13 would have been most visible. The best candidate image from the Voyager 1 encounter is 34740.42, but in this image the moonlet's predicted $\Sigma DN = 8$, which is less than that for any Voyager 2 frame in which 1981S13 was seen. Nevertheless, a very subtle bright spot is indeed visible within the gap, although the noise in the image means that this cannot be considered a definite detection. If it is assumed to be real, then the best-fit mean motion becomes $626.0463 \pm 0.0004^\circ$ per day. It seems, however,

FIG. 3 Sequence of five horizontal strips from Voyager 2 images 43904.59 to 43905.23, showing 1981S13 traversing the (relatively fixed) field of view. The time interval between frames is roughly 5 min. For each image a small box identifies the location of the moonlet, and the same region is shown enlarged at the upper left.



that a future detection of 1981S13 will be required to settle the question of whether it was truly detected by Voyager 1.

The moonlet wake revisited

Even without a definite detection by Voyager 1, 1981S13 now has the most accurately determined orbital location in Saturn's ring system. It is instructive to re-examine how this moonlet is related to the visible wake. Using the improved model for the geometry of the Voyager 2 PPS scan⁵ and formulae derived by Showalter *et al.*,² one can predict the positions of optical depth extrema in the wake and compare them with the measured locations (Fig. 5). But *a priori* one is still free to associate any predicted maximum with any observed maximum, so Fig. 5 shows the radial distance between each measured extremum and its predicted location, based on three different possible associations. The middle case, identified by $m_{\text{ref}} = 88$, is clearly the best fit as the calculated radial offset has a nearly constant value. The parameter m_{ref} is defined in ref. 2 and is constrained to be an even integer; it counts the number of extrema in the edge ripples between the occultation track and the moonlet. For comparison, the preferred value from the wake modelling alone was 90. This apparent misidentification, by a single wavelength, fully accounts for the discrepancy between the predicted and measured orbits of 1981S13.

In Fig. 5, the constant radial offset strongly suggests that the radial scale assumed for Saturn's rings⁵ is too large by 2.9 ± 0.8 km at the Encke gap. Such an offset is compatible with absolute uncertainties in the Voyager trajectory, which play a key part in determining the radial scale of the occultation profiles. The orbital radius of 1981S13 is derived from its mean motion and is therefore independent of trajectory errors; uncertainties arising from imprecise knowledge of Saturn's gravitational field should amount to ~ 0.1 km^{8,9}.

The constant radial offset also suggests that collective effects have not significantly altered this particular wake pattern, observed only 29.8° away from 1981S13. But the RSS wake pattern obtained $159.4^\circ \pm 1.6^\circ$ from the moonlet has been modified; the measured wavelength is systematically 10% larger than predicted. Nevertheless, in the outer wake of the PPS scan, taken a full 330.2° away in longitude, the wavelength is again very close to that predicted. This surprising result may indicate the differing influences of collisions and self-gravity on the evolution of a wake pattern. Self-gravity effects develop rapidly and increase the radial wavelength². Collisional effects grow more slowly with angular distance from the moonlet, but eventually dominate and tend to reduce the wavelength. This is qualitatively similar to what is observed; more-detailed models

of wake dynamics^{10,11} will ultimately have to account for these observations.

Discussion

Whereas a wake pattern represents the short-term disturbance in a ring from a single moonlet passage, the long-term effect of many such passages is to transfer angular momentum and thereby repel or 'shepherd' ring material¹². By this process, 1981S13 is undoubtedly responsible for keeping the Encke gap open: ring material is bounded at the point where its collisional spreading is balanced by the torque induced from each moonlet passage. The original shepherding theory assumed that a moonlet's perturbations are damped out in less than 360° of ring longitude, as is observed in the Encke wake. More recent models for the confinement of uranian and neptunian rings, as well as the outer edge of Saturn's B ring, have instead invoked isolated orbital resonances¹³⁻¹⁵. Although the process is still called 'shepherding', in this case the ring is bounded at a discrete location where the perturbations induced by repeated satellite passages add constructively. Hence, the Encke gap seems to be the best known example of the shepherding process as it was originally conceived.

Interestingly, the moonlet seems to fall within the radial limits of a narrow ringlet in the Encke gap (see Fig. 1). Dermott *et al.*¹⁶ were the first to note that ring material can be trapped in stable orbits sharing the same mean motion as a small moon; this confinement mechanism was originally proposed to explain the narrow uranian rings, as a rival to the two-moon shepherding model¹². It is ironic, therefore, that both models now seem to be illustrated simultaneously in Saturn's Encke gap. The Encke moonlet is capable of confining ring material to a zone ~ 30 km wide¹⁷, which readily encompasses the visible material.

The Voyager images reveal this ringlet to be far brighter in forward-scattered light than in backscatter, suggesting that it is composed primarily of micrometre sized dust. Although 1981S13 can ensure the orbital stability of this ring, such fine dust will be destroyed by micrometeoroid impacts over 10^6 – 10^7 years¹⁸. Hence, the material must be replenished from some nearby source, as is the case for the other faint, dusty planetary rings¹⁹. But this flux of micrometeoroids will also continually excavate new material from the surface of 1981S13 (along with any other macroscopic bodies in the vicinity), and this provides a plausible source for the dust. Burns *et al.*¹⁹ find that most of the particles ejected by high-velocity impacts from a soft planetary surface leave at velocities of ~ 1 m s⁻¹, which would disperse them over ~ 15 km in semimajor axis. This radial spread is small enough for most ejecta to remain trapped within the stable orbit zone

FIG. 4 Measured longitude of 1981S13 as a function of time for each of the Voyager 2 images in which it was seen. Longitudes are given relative to the original predicted orbital position; times are given relative to the moment of Voyager image 43905.11. The best straight-line fit is also shown.

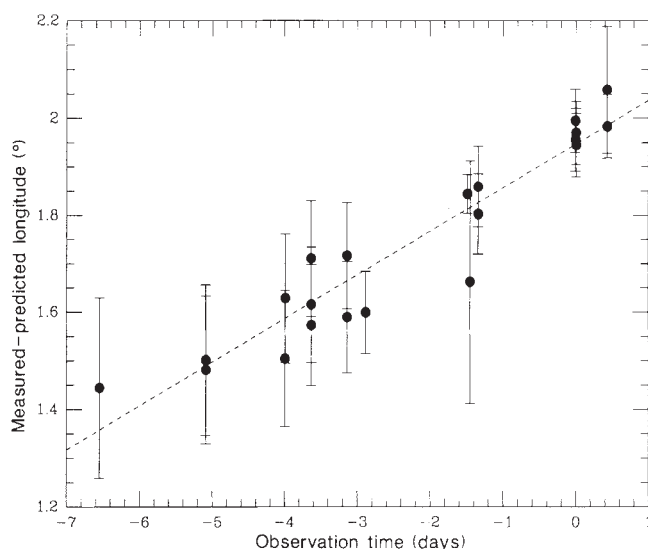
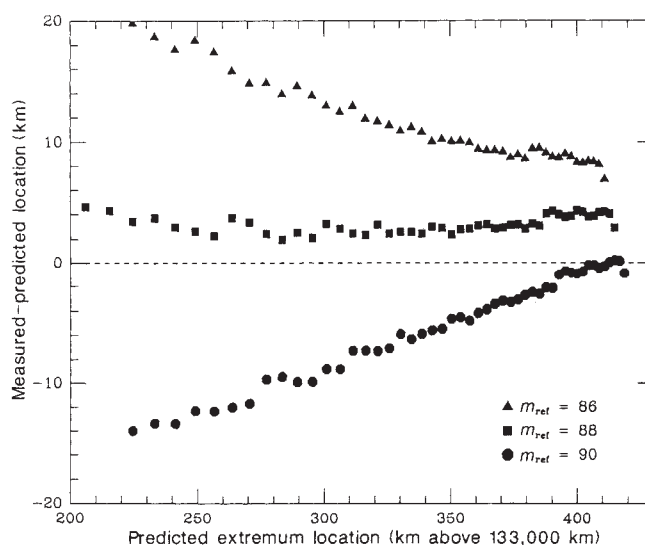


FIG. 5 Measured location of each extremum in the PPS wake compared with its predicted location, based on the best-fit orbit for 1981S13. Three different associations between measured and predicted extrema are shown, corresponding to whether the moonlet is 43, 44 or 45 edge ripples away ($m_{\text{ref}} = 86, 88$ or 90). The constant value of the central case strongly suggests that this is the correct association; it requires a 2.9-km reduction to the radial scale of the PPS profile.



of 1981S13. If this is the dominant mechanism at work, then the ring material will accumulate until it reaches a steady state, at which new dust is ejected from the moonlet's surface at the same rate that old dust is destroyed. This equilibrium point occurs when the ring cross-section matches that of the source moon multiplied by its impact yield factor Y , the ratio of mass ejected to mass striking the surface; perhaps surprisingly, it is independent of the micrometeoroid flux. Assuming a $Y \approx 10^4$ for impacts into the moonlet¹⁸, a ringlet of radial width 20 km will have an optical depth of ~ 0.1 , which is very compatible with the observations (Fig. 1).

In reality, the origin of the Encke ringlet cannot be as simple as this speculative model suggests. Collisions among any larger particles will augment the dust population, but this process is difficult to quantify without a more detailed understanding of the distribution of particle sizes in the ringlet. In addition, at least one other incomplete ringlet is visible in some Voyager images, and does not have the benefit of a large body like 1981S13 to confine or replenish its dust population. Furthermore, nothing in this model explains either ringlet's azimuthal variability, although the gravitational perturbations of 1981S13 are undoubtedly involved in some manner^{15,20}.

Beyond the Encke gap, 1981S13 does not seem to play an important part in Saturn's ring system. Its mass is $\sim 1\%$ that of the A ring from the Encke gap outward²¹, which is too small to affect the ring system's evolution. But its origin may be of greater interest, as the body clearly could not have accreted at its current location so far inside Saturn's Roche limit. If the moon and rings have a common origin, then the rings must also have been brought in from elsewhere, rather than having originated in something like their current form. Supposing that the rings arose from the breakup of one or more bodies in Saturn's distant past²², it is plausible that 1981S13 is merely the largest remaining fragment, and it might be better regarded as Saturn's largest ring particle rather than its smallest moon.

The detailed inspection of the outer planets by Voyagers 1 and 2 shows that rings and small satellites are commonly interspersed. It is therefore not surprising to find a moon embedded in the main rings of Saturn, although 1981S13 is the first such body to be identified; Voyager imaging sequences failed to reveal any comparable moonlets in the Cassini division²³. The significance of 1981S13 lies not so much in the fact that it is a new moon in the saturnian family, but in the diversity of interactions between rings and satellites that it helps to clarify. □

Received 8 April; accepted 24 May 1991.

1. Cuzzi, J. N. & Scargle, J. D. *Astrophys. J.* **292**, 276–290 (1985).
2. Showalter, M. R., Cuzzi, J. N., Marouf, E. A. & Esposito, L. W. *Icarus* **66**, 297–323 (1986).
3. Lane, A. L. *et al. Science* **215**, 537–543 (1982).
4. Marouf, E. A., Tyler, G. L. & Rosen, P. A. *Icarus* **68**, 120–166 (1986).
5. Nicholson, P. D., Cooke, M. L. & Pelton, E. *Astr. J.* **100**, 1339–1362 (1990).
6. Thomas, P., Veverka, J. & Dermott, S. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 802–835 (University of Arizona Press, Tucson, 1986).
7. Porco, C., Danielson, G. E., Goldreich, P., Holberg, J. B. & Lane, A. L. *Icarus* **60**, 17–28 (1984).
8. Nicholson, P. D. & Porco, C. C. *J. geophys. Res.* **93**, 10209–10224 (1988).
9. Campbell, J. K. & Anderson, J. D. *Astr. J.* **97**, 1485–1495 (1989).
10. Borderies, N., Goldreich, P. & Tremaine, S. *Icarus* **80**, 344–360 (1989).
11. Stewart, G. R. *Bull. Am. Astr. Soc.* **22**, 1045 (1990).
12. Goldreich, P. & Tremaine, S. *Nature* **277**, 97–99 (1979).
13. Borderies, N., Goldreich, P. & Tremaine, S. *Nature* **299**, 209–211 (1982).
14. Porco, C. C. & Goldreich, P. *Astr. J.* **93**, 724–729 (1987).

15. Porco, C. *Bull. Am. Astr. Soc.* **22**, 1043 (1990).
16. Dermott, S. F., Gold, T. G. & Sinclair, A. T. *Astr. J.* **84**, 1225–1234 (1979).
17. Dermott, S. F. & Murray, C. D. *Icarus* **48**, 1–11 (1981).
18. Durisen, R. H. *et al. Icarus* **80**, 136–166 (1989).
19. Burns, J. A., Showalter, M. R. & Morfill, G. E. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 200–272 (University of Arizona Press, Tucson, 1984).
20. Showalter, M. R. & Burns, J. A. *Icarus* **52**, 526–544 (1982).
21. Cuzzi, J. N. *et al. in Planetary Rings* (eds Greenberg, R. & Brahic, A.) 73–199 (University of Arizona Press, Tucson, 1984).
22. Harris, A. W. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 641–659 (University of Arizona Press, Tucson, 1984).
23. Smith, B. A. *et al. Science*, **215**, 504–537 (1982).

ACKNOWLEDGEMENTS. I thank J. Cuzzi for helpful discussions, P. Nicholson for assistance with the Voyager occultation geometry and gravitational field models, and E. Eliason, S. Kazz and L. Pieri of the Planetary Data Systems Imaging Node for their assistance in locating many of the Voyager images. This work was supported by NASA/Ames.

Cloned and expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase

David S. Bredt*, Paul M. Hwang*, Charles E. Glatt*, Charles Lowenstein*, Randall R. Reed*† & Solomon H. Snyder*‡

* Departments of Neuroscience, Pharmacology and Molecular Sciences, Psychiatry and Behavioral Sciences, Johns Hopkins Medical Institutions, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

† Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

Nitric oxide is a messenger molecule, mediating the effect of endothelium-derived relaxing factor in blood vessels and the cytotoxic actions of macrophages, and playing a part in neuronal communication in the brain. Cloning of a complementary DNA for brain nitric oxide synthase reveals recognition sites for NADPH, FAD, flavin mononucleotide and calmodulin as well as phosphorylation sites, indicating that the synthase is regulated by many different factors. The only known mammalian enzyme with close homology is cytochrome P-450 reductase.

NITRIC OXIDE (NO) has recently been discovered to be a messenger for blood vessels and neurons. It accounts for the activity of endothelium-derived relaxing factor (EDRF), which stimulates vasodilatation by releasing NO from the endothelium, which in turn acts on adjacent smooth muscle¹⁻³. NO is also responsible for the cytotoxic actions of macrophages⁴. In the brain, NO mediates the actions of the excitatory neurotransmitter glutamate in stimulating cyclic GMP concentrations, as inhibition of NO synthase (NOS; EC number 1.14.23) blocks elevations in cGMP^{5,6}. Immunohistochemical studies have localized NOS to particular neuronal populations in the brain and periphery⁷. Inhibitors of NOS block physiological relaxation of the intestine induced by neuronal stimulation, indicating that NO has the properties of a 'neurotransmitter'⁸⁻¹⁰. Thus, NO seems to be a novel type of neuronal messenger, in that, unlike conventional neurotransmitters, it is not stored in synaptic vesicles and does not act on typical receptor proteins of synaptic membranes. Its actions in activation of guanylyl cyclase¹¹ depend on its free-radical character, raising the question of whether other free-radical systems could be biological messengers.

NOS is a highly regulated enzyme. The brain-endothelial enzyme has an absolute requirement for calmodulin¹². NOS requires NADPH (refs 13, 14) and also has tightly bound flavins (D.S.B. and S.H.S., unpublished observations). Brain NOS is stoichiometrically phosphorylated by cyclic AMP-dependent protein kinase, protein kinase C and calcium/calmodulin-dependent protein kinase (our unpublished observations). NOS also has a recognition site for arginine, which is converted to NO and citrulline by a novel but so far uncharacterized process which requires molecular oxygen¹⁵.

So that we can find out how NOS catalyses the transformation of arginine to NO, how this process is regulated by factors, and whether NOS resembles other enzymes that would give rise to

free radicals, we have determined its full amino-acid sequence. We report here the cloning and expression of a complementary DNA for brain NOS. The sequence of NOS has striking similarities to that of cytochrome P-450 reductase, another oxidative enzyme which is also unusual in having recognition sites for two flavins and NADPH (ref. 16).

Cloning of NOS

NOS protein was purified from rat cerebellum as previously described¹², digested with trypsin and purified fragments sequenced. Because of the low abundance of NOS protein and messenger RNA, we developed a sensitive two-phased polymerase chain reaction (PCR) cloning strategy which should be useful for cloning other rare mRNA molecules (see legend to Fig. 1). This yielded a fragment of 599 base pairs (bp) which had the expected amino acids next to each of the primers and was used to screen 10⁶ clones of a rat brain cDNA library (Stratagene), and eight independent recombinant phage were isolated. Three of these clones, spanning 5,057 bases, had an open reading frame of 4,287 bases encoding a protein of 1,429 amino acids and of relative molecular mass 160,458 ($M_r \sim 160K$) and incorporates all 21 peptides that had been sequenced previously. The start codon was assigned to the first methionine in the open reading frame at nucleotide 349 because the neighbouring bases, ATACC(ATG)G, conform well to the consensus for initiation of translation¹⁷.

Functional expression of NOS

The full-length coding cDNA was inserted into an expression vector using a cytomegalovirus promoter which was transfected into human kidney 293 cells. Coomassie-blue staining of cell extracts collected 3 days after transfection reveals one band in transfected cells not present in untransfected cells and having an M_r of 160K, corresponding to the size of purified NOS (ref. 12; Fig. 2). NOS protein corresponds to about 1-2% of total protein in the transfected cells, whereas negligible amounts are present in the untransfected cells. Western blot analysis using an antiserum that is highly selective for NOS (ref. 7), reveals a single 160K band in the transfected cells which is not present in the untransfected cells and which corresponds to NOS in cerebellar extracts (Fig. 2). The immunoreactive band identified from 30 μ g of transfected cells is more intensely labelled than is the band from 200 μ g of cerebellar tissue, consistent with a 10-fold enrichment of NOS in the transfected cells over the cerebellum.

NOS catalytic activity in transfected cells was assayed in three ways, measuring (1) the conversion of [³H]arginine to [³H]citrulline; (2) the production of nitrite from unlabelled arginine; and (3) the enhancement of endogenous guanylyl cyclase activity in response to newly synthesized NO (Fig. 3). With all three assays there is negligible NOS activity in untransfected cells, whereas there is great activity in transfected cells. This activity is about 10 times greater than that in cerebellar

‡ To whom correspondence should be addressed.

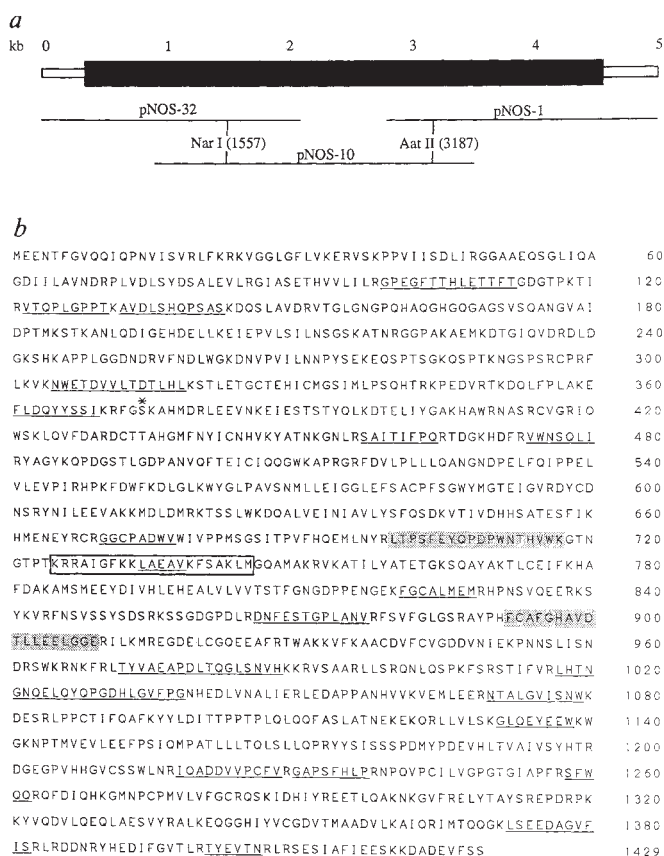


FIG. 1 Amino-acid sequence of NOS (single-letter amino-acid notation). **a**, Diagram of the overlapping partial cDNAs encoding NOS and the restriction sites used for their ligation. The translated sequence is indicated by a solid bar. **b**, Predicted amino-acid sequence of NOS. The tryptic peptides sequenced from the purified protein are underlined. The two peptides used for the initial PCR cloning are shaded. Asterisk indicates the consensus sequence for phosphorylation by protein kinase A; the predicted binding site for calmodulin is boxed.

METHODS. The two-round PCR cloning strategy is as follows. First, PCR primers were designed on the basis of the amino-acid sequences of the two largest peptides sequenced, 17 and 18 amino acids, respectively. For each of these peptides, a pair of fully degenerate oligonucleotides were constructed to match the six amino acids at the N and C termini, and these were used to amplify the short intervening sequences between the primers by PCR at 42 °C annealing temperature on first-strand complementary DNA derived from rat brain RNA. After amplification, the single prominent PCR product of the expected size from each reaction was excised from an agarose gel, cloned and sequenced to reveal the nucleotides encoding the five or six central amino acids of each peptide. Nondegenerate oligonucleotides corresponding to the sequences determined from these central regions were used in a second round of PCR done at a much higher annealing temperature (63 °C) on first-strand rat brain cDNA. Two pairs of oligonucleotides were synthesized as the relative order of the two peptides in the protein sequence of NOS were unknown. One pair of oligonucleotides gave no PCR product, whereas the other produced a single 599-bp product. This technique which was used for plaque hybridization to isolate cDNA clones from a rat brain library. The PCR product was random-prime labelled (1×10^6 c.p.m. ^{32}P ng $^{-1}$), hybridized to replicate filters overnight and the filters were washed with high stringency ($0.1 \times \text{SSC}$ medium, 0.1% SDS, 65 °C). Overlapping clones were sequenced by the dideoxy chain termination method using a commercial kit, Sequenase (USB). All PCR primers included nine-bp sequences at the 5' end which allowed subcloning into *Eco*R1 or *Bam*H1 sites. The first round of PCR (42 °C annealing temperature, 36 cycles) was performed with the following primers: peptide 1, CCGGAATTCCTNACNCCNT-NTT(T/C)GA and GCCGGATCC(C/T)TTNAC(A/G)TGNGT(A/G)TTCCA; peptide 2, CCGGAATTCCTT(T/C)TG(T/C)GCNTT(T/C)GGNCA, and GCCGGATCC(T/C)TCNCCNCCNAG(T/C)TC(T/C)TC. Primers used for the second round of PCR (63 °C annealing temperature, 36 cycles) were CCGGAATTCGAATAC-CAGCCTGATCCATGGAA and GCCGGATCCTCCAGGAGGTGTCCACCGCATG. The start codon is at position 349 and the stop at 4,636.

extracts, the richest known mammalian source of NOS activity of the brain-endothelial type. The NOS activity is about 1% of that of purified brain NOS, suggesting that it accounts for 1% of the protein in the transfected cells, which fits with the estimate based on Coomassie-staining of extracts from these cells. NOS activity of the transfected cells corresponds to the properties of the enzyme characterized in the brain¹². Enzyme activity is virtually lost in the absence of NADPH or calcium and is inhibited by trifluoperazine, a calmodulin antagonist; elevations in guanylyl cyclase activity are blocked by haemoglobin, which binds NO (Fig. 3).

NOS mRNA expression

To evaluate the expression of the NOS gene, we analysed various tissues and brain regions by northern blotting (Fig. 4). There is a prominent 10.5-kb band in total cerebellar RNA. Because the coding sequence for NOS is only about 4.3 kb, around half of its mRNA is not translated. We do not detect NOS mRNA in kidney, liver, skeletal muscle, stomach or heart. In the brain, NOS mRNA levels are highest in the cerebellum, followed in descending order by olfactory bulb, colliculi, hippocampus and cerebral cortex. The relative amounts of mRNA in these brain regions correspond closely to the relative amounts of NOS protein and catalytic activity⁷.

We localized NOS mRNA *in situ* at a microscopic level using hybridization (Fig. 5a). It is most prominent in the cerebellum, olfactory bulb, and pedunculo-pontine tegmental nucleus, structures which are seen to be highly enriched in NOS when using immunohistochemistry (ref. 7; and unpublished observations). In the cerebellum, NOS mRNA is most enriched in the granule cell layer, and also enriched in the main olfactory bulb, dentate gyrus of the hippocampus, supraoptic nucleus and superior and

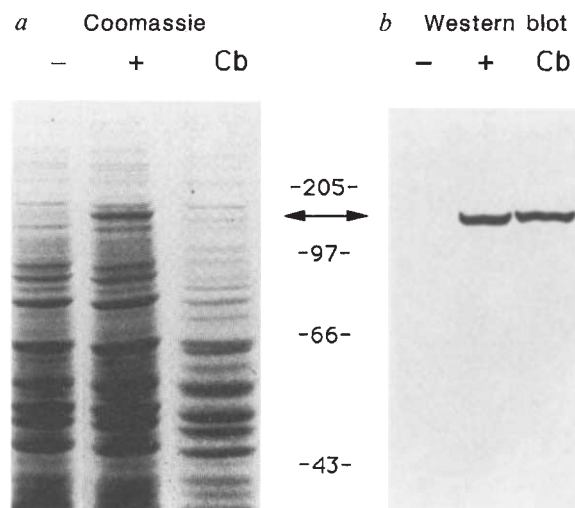


FIG. 2 Expression of NOS protein. **a**, Coomassie-blue staining and **b**, western blot analysis of soluble extracts from rat cerebellum (Cb) and human kidney cells following transfection with control expression vector (-) or NOS vector (+). Positions of M_r markers are shown ($M_r \times 10^{-3}$); double-headed arrow at 160K indicates the position of NOS protein.

METHODS. cDNA clones spanning the full open reading frame for NOS were combined as shown in Fig. 1a, excised from bluescript with *Cla*I and *Xba*I, and ligated into the expression vector p-CIS-2 (Genentech). Human 293 kidney cells were transfected with calcium phosphate as described⁴², cells were collected 3 days after transfection, homogenized in buffer containing 1 mM EDTA, and centrifuged at 100,000g for 60 min. Soluble extracts from transfected kidney cells (30 μg per lane) or rat cerebellum (30 μg for Coomassie-blue staining, 200 μg for western blot) were separated on a 7.5% SDS polyacrylamide gel, then either stained with Coomassie blue or transferred to nitrocellulose, probed with anti-NOS antibody⁷ overnight at 1:5,000 dilution and the bands visualized with an alkaline phosphatase-linked secondary antibody.

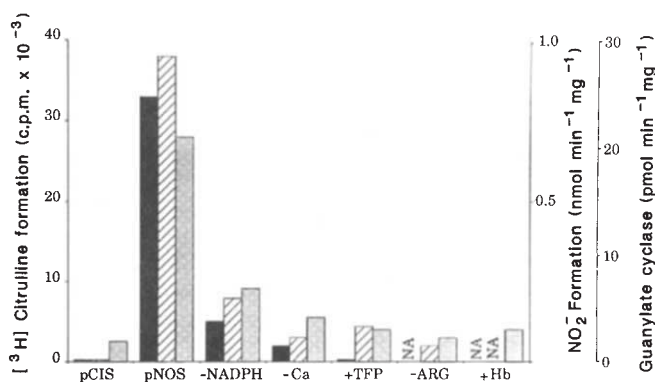


FIG. 3 Functional expression of NOS enzyme activity. NOS activity was measured in control transfected kidney cells (pCIS) and in cells transfected with NOS expression vector (pNOS). Enzyme activity was assayed in three ways: (1) conversion of [³H]arginine to [³H]citrulline (solid bars); (2) formation of NO₂⁻ from unlabelled arginine (striped bars); and (3) activation of endogenous soluble guanylyl cyclase by arginine (dotted bars). For the first two sets of bars, enzyme activity was measured with complete reaction mixtures and found only in NOS-transfected cells. The final five sets all use NOS-transfected tissue. Data are means of triplicate determinations from a representative experiment.

METHODS. Kidney cells were transfected, soluble extracts were prepared as described in Fig. 2 and dialysed for 12 h to remove most endogenous arginine and NADPH. [³H]Citrulline formation from [³H]arginine and NO₂⁻ production from unlabelled arginine were monitored as described⁵. Guanylyl cyclase activity was assayed by monitoring cyclization of [α -³²P]GTP to [³²P]cGMP as described³⁸. The standard reaction mixture was 20 mM Tris-HCl (pH 7.4), 1 μ M free Ca²⁺, 1 mM NADPH. For [³H]citrulline assays, 1 μ M [³H]arginine (100,000 c.p.m.) was added; for nitrate assays, 1 mM arginine was added; for guanylyl cyclase assays, 1 mM arginine, 1 mM cGMP, 0.1 mM α -[³²P]GTP, 5 mM Mg²⁺, 1 mM isobutylmethyl xanthine and 50 units ml⁻¹ superoxide dismutase were added. TFP, 100 μ M trifluoperazine; Hb, 10 μ M haemoglobin; NA, not applicable.

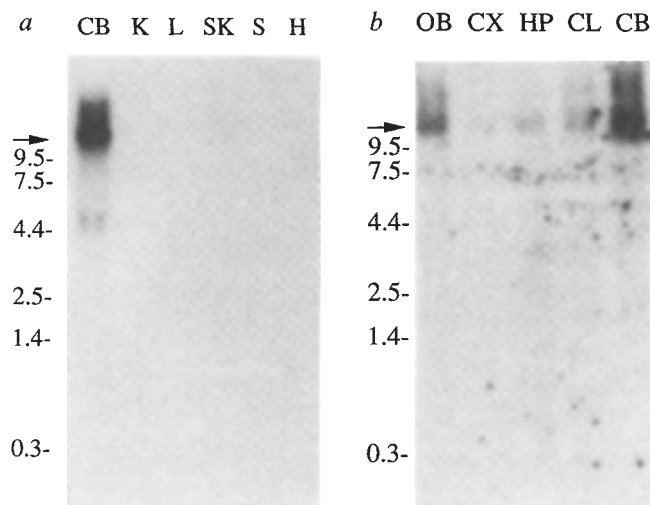


FIG. 4 NOS mRNA expression in brain and peripheral tissues. *a*, Northern blot analysis of total RNA (20 μ g) from cerebellum (CB), kidney (K), liver (L), skeletal muscle (SK), stomach (S) and heart (H). The arrow indicates the NOS message present only in cerebellum at 10.5 kb. *b*, Regional distribution of NOS mRNA in brain reveals highest levels in cerebellum (CB), followed by olfactory bulb (OB), colliculi (CL), hippocampus (HP) and cerebral cortex (CX). Positions of *M_r* (K) markers are shown.

METHODS. RNA was purified from rat tissues by tissue lysis in guanidinium isothiocyanate followed by centrifugation through CsCl (ref. 44). The RNA samples were separated on a formaldehyde agarose gel, probed with a random prime-labelled (3×10^6 c.p.m. ³²P ng⁻¹; 10^8 c.p.m. per filter) full-coding region cDNA (bases 1–5,057), washed at high stringency (0.5×SSC medium, 0.1% SDS, 65 °C), and exposed to X-ray film for 42 days.

inferior colliculi. The white dots in the cerebral cortex, caudate putamen and basal forebrain represent isolated NOS-containing neurons, the distribution of which resembles that seen by immunohistochemistry¹⁸. High-power micrographs of the caudate putamen reveal NOS mRNA localized to large aspiny neurons (Fig. 5c), which are also selectively enriched in NADPH diaphorase staining (Fig. 5b).

Primary structure of NOS

The NOS amino-acid sequence has a number of recognition sites needed for enzyme's function (Figs 1 and 6). There is a basic amphipathic α helix calmodulin-binding consensus

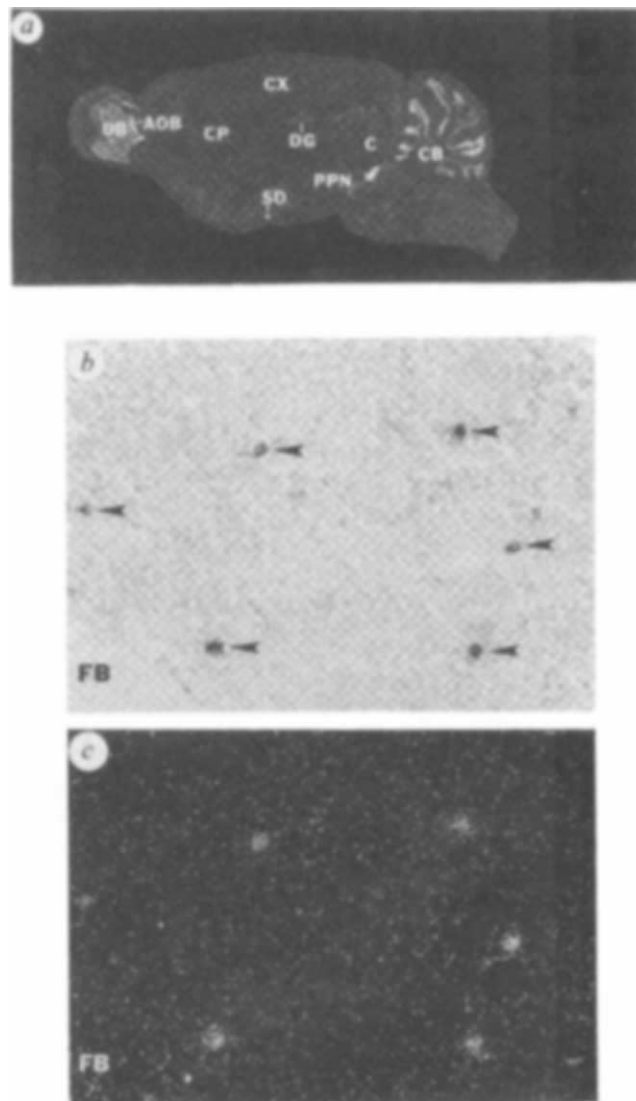


FIG. 5 Localization of NOS mRNA in brain. *a*, *In situ*-hybridization to NOS mRNA in a sagittal section of rat brain in darkfield shows densest message levels in granule cell layer of cerebellum (CB), accessory olfactory bulb (AOB) and pedunculopontine tegmental nucleus (PPN). High densities of grains are also apparent in the main olfactory bulb (OB), dentate gyrus (DG), supraoptic nucleus (SO) and superior and inferior colliculi (C), white dots scattered throughout the cerebral cortex (CX), caudate putamen (CP) and basal forebrain represent isolated NOS-containing neurons in these regions. *b*, High-power lightfield micrograph showing coincidence of NADPH diaphorase staining of aspiny neurons (arrowheads) in caudate putamen with *c*, NOS mRNA *in situ* on the same section viewed under darkfield. FB, fibre bundle.

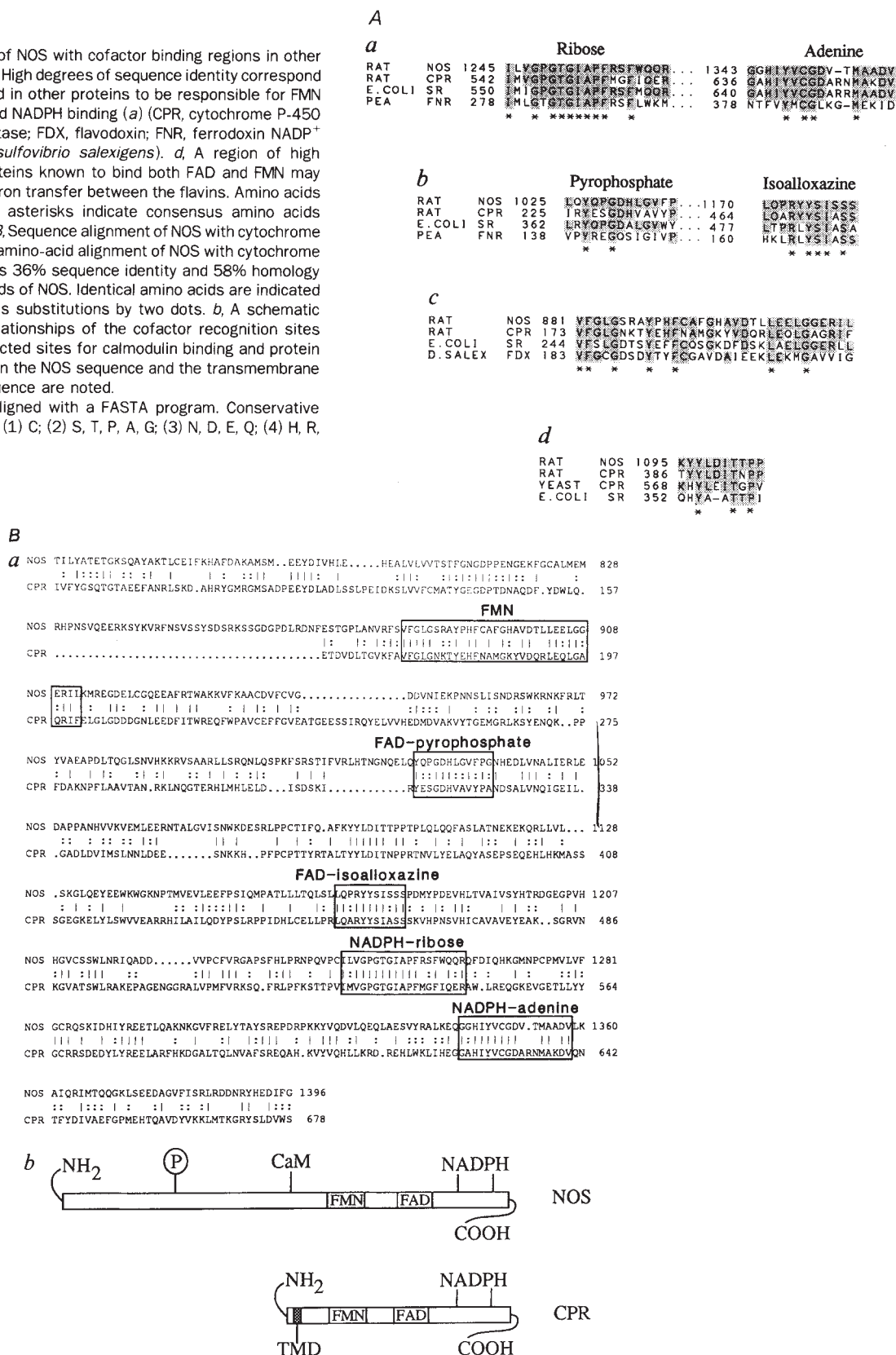
METHODS. *In situ* hybridization was done as described⁴⁵ with a pool of three end-labelled 45-base oligonucleotides complementary to bases 211–266, 795–840 and 4,712–4,757. For the colocalization, NADPH diaphorase staining was done immediately before *in situ* hybridization as described⁴⁶.

sequence¹⁹ at position 725–745, consistent with its requirement of calmodulin¹². A cAMP-dependent protein kinase phosphorylation consensus sequence²⁰, KRFGS, is at position 473, consistent with a serine phosphorylation by cAMP-dependent protein kinase and fitting with the observed phosphorylation by this enzyme. There are no consensus sequences for phosphorylation by protein kinase C or calcium/calmodulin protein kinase

which are selective enough to be identified. An NADPH-binding domain makes up the area from amino acids 1,204 to 1,429 (ref. 21). Well-defined sites representing the point of contact with the ribose and adenine rings are at amino acids 1,245–1,263 and 1,343–1,358, respectively, and are closely homologous to similar sites in cytochrome P-450 reductase, sulphite reductase and ferredoxin NADP⁺ reductase^{22–26} (Fig. 6a).

FIG. 6 A sequence alignment of NOS with cofactor binding regions in other electron transferase enzymes. High degrees of sequence identity correspond to regions previously identified in other proteins to be responsible for FMN binding (c), FAD binding (b) and NADPH binding (a) (CPR, cytochrome P-450 reductase; SR, sulphite reductase; FDX, flavodoxin; FNR, ferredoxin NADP⁺ reductase; and D. Salex, *Desulfovibrio salexigens*). d, A region of high homology present only in proteins known to bind both FAD and FMN may be involved in facilitating electron transfer between the flavins. Amino acids identical to NOS are shaded; asterisks indicate consensus amino acids present in all four sequences. B, Sequence alignment of NOS with cytochrome P-450 reductase (CPR). The predicted sites for calmodulin binding and protein kinase A phosphorylation within the NOS sequence and the transmembrane domain (TMD) in the CPR sequence are noted.

METHODS. Sequences were aligned with a FASTA program. Conservative substitutions were defined as (1) C; (2) S, T, P, A, G; (3) N, D, E, Q; (4) H, R, K; (5) M, I, L, V; (6) F, Y, W.



Enzymes requiring NADPH typically use other cofactors as well. Tetrahydrobiopterin enhances the activity of macrophage NOS (refs 27, 28) and stimulates porcine brain NOS (ref. 29). The brain NOS sequence has consensus binding sites for the flavins FMN and FAD (Fig. 5). There are distinct binding sites for the pyrophosphate and isoalloxazine moieties at the FAD region²¹. These flavin-binding sites are specifically conserved in a family of electron transfer proteins²¹. Using fluorescence spectroscopy we have demonstrated the stoichiometric occurrence of tightly bound FMN and FAD in purified rat brain NOS (unpublished results).

The amino-acid sequence of NOS was compared with that of all sequenced proteins in GeneBank using a FASTA program. The closest homology is to rat cytochrome P-450 reductase (Fig. 6b). The two proteins are similar at the C-terminal half of NOS which, over 641 amino acids, has 36% identity and 58% close homology to cytochrome P-450 reductase. The only other protein with substantial homology to NOS is sulphite reductase which has 32% identity and 49% close homology to NOS over the C-terminal 641 amino acids of NOS. NOS, cytochrome P-450 reductase and sulphite reductase are a unique class of electron transferase enzymes characterized by binding sites for NADPH, FMN and FAD in the same polypeptide. For the two reductases, the flavins have been shown to make up a transport chain with electrons shuttling between the isoalloxazine rings. Specific sequences in cytochrome P-450 reductase might be responsible for coupling the FAD and FMN regions²¹. NOS has a similar sequence (Fig. 6a) and thus electron transfer between the flavins probably has a role in the mechanism of NO synthesis. The N-terminal half of NOS has no significant homology to any known protein.

NOS gene family

Clearly there are several distinct NOS enzymes. The brain and endothelium seem to use the same or similar calmodulin-dependent enzyme, brain NOS (bNOS), as both are recognized by a specific antiserum to bNOS (ref. 7). The macrophage³⁰ and neutrophil³¹ must each have distinct enzymes as neither requires calmodulin for activity and neither is labelled by antiserum to bNOS (ref. 7). We have recently identified a gene product highly homologous to bNOS that is expressed selectively in activated macrophages and is probably the macrophage enzyme, macNOS (C.L., D.S.B. and S.H.S., manuscript in preparation).

Discussion

Because of the selective close similarity of amino-acid sequence between cytochrome P-450 reductase and NOS, we wondered whether the two enzymes have analogous functions. Cytochrome P-450 reductase provides the electrons for the activities of the microsomal P-450 enzymes³². Histochemical analysis in the brain shows selective neuronal localizations for a number of P-450 enzymes^{33,34}, as well as for cytochrome P-450 reductase in discrete populations of catecholamine neurons³⁵. Besides these catecholamine neurons, cytochrome P-450 reductase also occurs in a number of other selected neuronal populations such as Purkinje cells of the cerebellum (M. Fotuhi, T. Dawson, S.H.S., unpublished results). In the liver and certain other peripheral tissues, the P-450 enzymes are primarily associated with drug metabolism. This would seem unlikely for P-450 enzymes and cytochrome P-450 reductase in the brain, especially in light of the localization of these enzymes to very discrete neuronal populations in the brain. Conceivably, oxidative products formed by the P-450 system have a messenger role in discrete neuronal populations similar to that of NO.

The reasons for selectively high production of NO by very particular populations of neurons have been obscure. Although NO mediates the effects of glutamate in elevating cGMP amounts in the cerebellum^{5,6}, the localization of NOS in brain does not resemble that of guanylyl cyclase mapped by immunohistochemistry³⁶, suggesting that NO has multiple functions in the brain. We find the mRNA for NOS to be colocalized with neurons selectively enriched in NADPH diaphorase, complementing our findings by immunohistochemistry, which demonstrate that NOS activity fully accounts for NADPH diaphorase activity^{18,37}. NADPH diaphorase has been characterized by histochemical staining to have an oxidative activity dependent on NADPH but not NADH³⁸. NADPH diaphorase neurons are selectively resistant to ischaemic destruction and survive effects of excitatory neurotoxins as well as the degenerative process of Huntington's disease³⁹⁻⁴². Thus, one function of NO may be to protect neurons from ischaemic and neurotoxic insults. NOS inhibitors selectively prevent glutamatergic neurotoxicity mediated by *N*-methyl-D-aspartate receptors in primary brain cultures, indicating that NO secreted by NOS neurons may kill adjacent neurons⁴³. □

Received 28 March; accepted 13 May 1991.

- Moncada, S., Palmer, R. M. J. & Higgs, E. A. *Biochem. Pharmacol.* **38**, 1709-1713 (1989).
- Furchgott, R. F. & Vanhoutte, P. M. *FASEB J.* **3**, 2007-2018 (1988).
- Ignarro, L. J. A. *Rev. Pharmacol. Tox.* **30**, 535-560 (1990).
- Hibbs, J. B. Jr., Vavrin, Z. & Taintor, R. R. *J. Immun.* **138**, 550-565 (1987).
- Bredt, D. S. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9030-9033 (1989).
- Garthwaite, J., Garthwaite, G., Palmer, R. M. J. & Moncada, S. *Eur. J. Pharmacol.* **172**, 413-416 (1989).
- Bredt, D. S., Hwang, P. M. & Snyder, S. H. *Nature* **347**, 768-770 (1990).
- Bult, H. *et al. Nature* **345**, 346-347 (1990).
- Gillespie, J. S., Liu, X. & Martin, W. *Br. J. Pharmacol.* **98**, 1080-1082 (1989).
- Ramagopal, M. W. & Leighton, H. J. *Eur. J. Pharmacol.* **174**, 297-299 (1989).
- Arnold, W. P., Mittal, C. K., Katsuki, S. & Murad, F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3203-3207 (1977).
- Bredt, D. S. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **87**, 682-685 (1990).
- Marletta, M. A., Yoon, P. S., Iyengar, R., Leaf, C. D. & Wishnok, J. S. *Biochemistry* **27**, 8706-8711 (1988).
- Palmer, R. J. M. & Moncada, S. *Biochem. biophys. Res. Commun.* **158**, 348-352 (1989).
- Kwon, N. S. *et al. J. Biol. Chem.* **265**, 13442-13445 (1990).
- Iyanagi, T. & Mason, H. S. *Biochemistry* **12**, 2297-2308 (1973).
- Kozak, M. *Nucleic Acids Res.* **15**, 8125-8148 (1987).
- Snyder, S. H. & Bredt, D. S. *Trends pharmacol. Sci.* **12**, 125-128 (1991).
- O'Neil, K. T. & DeGrado, W. F. *Trends biol. Sci.* **15**, 59-64 (1990).
- Kemp, B. E. & Pearson, R. B. *Trends biol. Sci.* **15**, 342-346 (1990).
- Porter, T. D. & Kasper, C. B. *Biochemistry* **25**, 1682-1687 (1986).
- Porter, T. D. & Kasper, C. B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 973-977 (1985).
- Ostrowski, J. *et al. J. Biol. Chem.* **27**, 15796-15808 (1989).
- Newman, B. J. & Gray, J. C. *Pl. molec. Biol.* **10**, 511-520 (1988).
- Helms, L. F., King, G. D. & Swenson, R. P. *Biochem. biophys. Res. Commun.* **168**, 809-817 (1990).
- Yabusaki, Y., Murakami, H. & Ohkawa, H. *J. Biochem.* **103**, 1004-1010 (1988).
- Tayeh, M. A. & Marletta, M. A. *J. Biol. Chem.* **264**, 19654-19658 (1989).
- Kwon, N. S., Nathan, C. F. & Stuehr, D. J. *J. Biol. Chem.* **264**, 20496-20501 (1989).

- Mayer, B., John, M. & Bohme, E. *FEBS Lett.* **277**, 215-219 (1990).
- Nathan, C. F. & Stuehr, D. J. *J. natn. Cancer Inst.* **82**, 726-728 (1990).
- Yui, Y. *et al. J. Biol. Chem.* **266**, 3369-3371 (1991).
- Lu, A. Y. H., Junk, K. W. & Coon, M. J. *J. Biol. Chem.* **244**, 3714-3721 (1969).
- Kapitulnik, J., Gelboin, H. V., Guengerich, F. P. & Jacobowitz, D. M. *Neuroscience* **20**, 829-833 (1987).
- Warner, M., Kohler, C., Hansson, T. & Gustafsson, J. A. *J. Neurochem.* **50**, 1057-1065 (1988).
- Haglund, L., Kohler, C., Haaparanta, T., Goldstein, M. & Gustafsson, J. *Nature* **307**, 259-262 (1984).
- Nakane, M., Ichikawa, M. & Deguchi, T. *Brain Res.* **273**, 9-15 (1983).
- Dawson, T. M., Bredt, D. S., Fotuhi, M., Hwang, P. M. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Thomas, E. & Pearce, A. G. E. *Acta Neuropath.* **3**, 238-249 (1964).
- Beal, M. F. *et al. Nature* **321**, 168-171 (1989).
- Koh, J.-Y., Peters, S. & Choi, D. W. *Science* **234**, 73-76 (1986).
- Ferrante, R. J. *et al. Science* **230**, 561-563 (1985).
- Vemura, Y., Kowall, N. W. & Beal, M. F. *Ann. Neurol.* **27**, 620-625 (1990).
- Dawson, V. L., Dawson, T. M., London, E. D., Bredt, D. S. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Pritchett, D. B. *et al. Science* **242**, 1306-1308 (1988).
- Schultz, G. & Bohme, E. *Methods of Enzymatic Analysis* 3rd edn Vol IV (eds Bergmeyer, H. V., Bergmeyer, J. & Grabl, M.) 379-389 (Verlag Chemie, Deerbeach, Florida, 1983).
- Davis, L. G., Dibrner, M. D. & Batley, J. F. *Basic Methods in Molecular Biology*, 130-135 (Elsevier, New York, 1986).
- Ross, C. A., MacCumber, M. W., Glatt, C. A. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2923-2927 (1989).
- Mizukawa, K., Vincent, S. R., McGeer, P. L. & McGeer, E. G. *J. comp. Neurol.* **279**, 281-311 (1989).

ACKNOWLEDGEMENTS. We thank P. Shenbagamurthi and V. Vogel for peptide sequencing, G. Pitt and L. Levin for advice and assistance with cDNA cloning and expression, A. Verma for helpful discussions and Genentech for the mammalian expression vector. This work was supported by the USPHS, the PSA, the International Life Sciences Institute and Bristol-Myers-Squibb.

A high-redshift IRAS galaxy with huge luminosity—hidden quasar or protogalaxy?

M. Rowan-Robinson*, T. Broadhurst*, A. Lawrence†, R. G. McMahon‡, C. J. Lonsdale§, S. J. Oliver*, A. N. Taylor*, P. B. Hacking§, T. Conrow§, W. Saunders||, R. S. Ellis¶, G. P. Efstathiou|| & J. J. Condon#

* School of Mathematical Sciences, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

† Department of Physics, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

‡ Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

§ Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, California 91125, USA and Jet Propulsion Laboratory

|| Department of Astrophysics, University of Oxford, 1 Keble Road, Oxford OX1 3RH, UK

¶ Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

National Radio Astronomy Observatory, Edgemont Road, Charlottesville, Virginia 22903, USA

DURING a survey intended to measure redshifts for 1,400 galaxies identified with faint sources detected by the Infrared Astronomy Satellite, we found an emission-line galaxy at a redshift of 2.286, and with the enormous far-infrared luminosity of 3×10^{14} times that of the Sun ($L_{\odot}$). The spectrum is very unusual, showing lines of high excitation but with very weak Lyman- α emission. A self-absorbed synchrotron model for the infrared energy distribution cannot be ruled out, but a thermal origin seems more plausible. A radio-quiet quasar embedded in a very dusty galaxy could account for the infrared emission, as might a starburst embedded in $1\text{--}10 \times 10^9 M_{\odot}$ of dust. The latter case demands so much dust that the object would probably be a massive galaxy in the process of formation. In either case, this is a remarkable object, and the presence of a large amount of dust in an object of such high redshift implies the generation of heavy elements at an early cosmological epoch.

On the release of the Infrared Astronomy Satellite (IRAS) Faint Source Survey (FSS), we embarked on a redshift survey of a large sample (1,400) of galaxies identified with FSS sources brighter than 0.2 Jy at 60 μm , covering an area of 700 square degrees¹. We have obtained redshifts for 95% of the sample. The remarkable object discussed here was discovered on the final night of the campaign.

IRAS Faint Source 10214+4724 lies at the limit of the flux range we selected for our survey, with a flux at 60 μm , $S_{60} = 0.205$ Jy. The flux density at 100 μm is $S_{100} = 0.57$ Jy, so the colour ratio, S_{100}/S_{60} , is rather cool. Analysis of the raw IRAS data at IPAC confirmed these fluxes (see Table 1) and yielded a marginal detection at 25 μm , 0.075 ± 0.031 Jy. It also gave a slightly improved position for the source.

Figure 1 shows a CCD (charge-coupled device) image centred on the original FSS position of the source, taken with the Hale telescope on 27 November 1990 by P. McCarthy, with the IRAS 1σ error ellipse superposed. We have taken spectra at various times of candidates A, B, C, G and F, on the Isaac Newton and William Herschel Telescopes on La Palma. All have spectra that are stellar or show no emission lines, except for object F, which has strong emission lines. (The discovery spectrum of object F, which was not visible on the Palomar Sky Survey plates, was obtained by great good fortune on the same frame as an observation of object C.) The emission lines would normally ensure that object F is the correct identification, and the unusual nature of the identified object makes coincidence unlikely, but we

confirmed the identification by taking a map of the field at 4.86 GHz with the C configuration of the Very Large Array (VLA) on 1 December 1990. By far the strongest source in the field is within one arcsec of the optical centre of object F. The optical image of object F (taken in seeing of 1.2–1.5") shows an unresolved core and an asymmetric diffuse extension 4" in extent. Table 1 summarizes the infrared optical and radio data on 10214+4724.

Figure 2 shows the result of adding two optical spectra of object F, taken with the Durham-Royal Greenwich Observatory Faint Object Spectrograph on the William Herschel Telescope on 15 and 19 May 1990. We identified six emission lines, giving a firm redshift of $z = 2.286$. Three further spectra, taken on 11 March 1991, extended the observations down to 3,500 Å, showing four further lines: Ly α (1216 Å), N v (1,240 Å), O iv + Si iv (1,400 Å), and N v (1,484 Å). Remarkably, Lyman- α is weaker than N v. The emission-line spectrum is very different from that of quasars, in particular showing no broad lines. It is also very different from nearby starburst galaxies. It is similar to those seen in Seyfert 2 galaxies, liners and high redshift ($z > 1$) radiogalaxies^{2–4}, except that Ly α is exceptionally weak (in the radiogalaxies Ly α is typically 30 times stronger than the 1,240 Å N v line). Also the radio flux is several orders of magnitude smaller than that of such radiogalaxies. The emission-line spectrum is similar to that seen in an oxygen-rich supernova remnant in the Small Magellanic Cloud⁵, except that the 1,640 Å He II line is absent in this supernova remnant, and to the Cygnus Loop⁶, except that the 2,424 Å [Ne IV] line is absent in the latter.

The rest-frame luminosity at 6–36 μm is $3 \times 10^{14} L_{\odot}$ ($H_{\odot} = 50$, $\Omega_{\odot} = 1$), 10 times more luminous than the most luminous IRAS galaxy previously known, the Seyfert 2 galaxy 09104+4109 (ref. 7). In the rest frame of 10214+4724, the R-magnitude corresponds to a luminosity of $7.4 \times 10^{11} L_{\odot}$ at 2,130 Å. A service

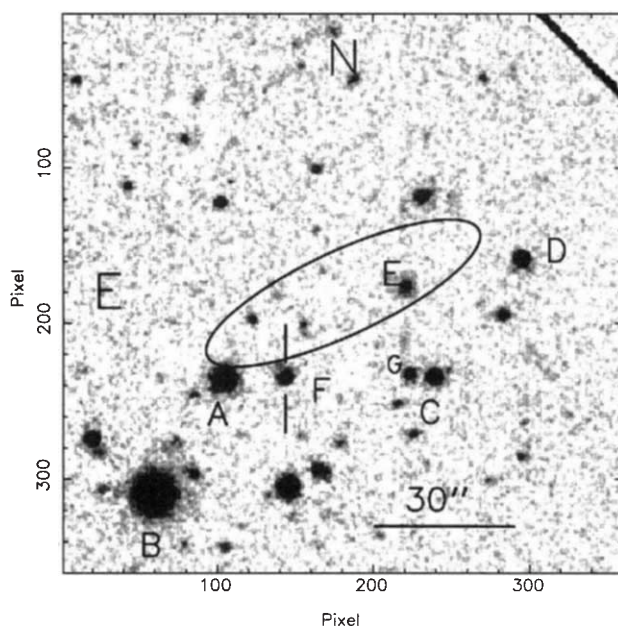


FIG. 1 A 300-s CCD image, taken in the Gunn r -band (6,550 Å) on the Hale telescope, of the region centred on the original FSS position of 10214+4724, showing location of objects A–G, together with error ellipse for IRAS source. The Addscan position is 20 arcsec south of this. The VLA source coincides with object F, the galaxy with $z = 2.286$. Objects A, B, C and G have spectra that are stellar or have no emission lines. Object E is stellar, was not detected by the VLA and is one magnitude fainter than F, so is unlikely to be the correct identification. Object F consists of a compact unresolved (in 1.5" seeing) component with a faint low surface brightness (25–26 mag per sq arcsec) asymmetric extension of total extent 4". The latter comprises 10–20% of the total optical flux.

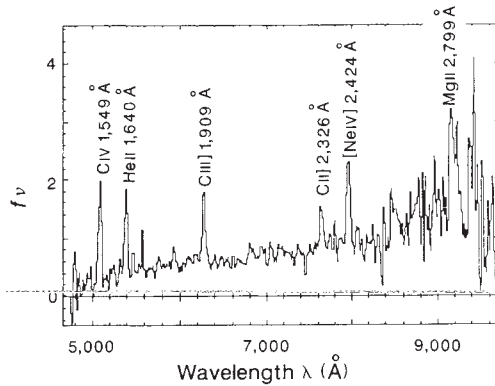


FIG. 2 Spectrum of 10214+4724F, with emission lines identified and labelled with their rest wavelengths. The spectrum has been approximately calibrated and the atmospheric A band has been removed using observations of star A.

observation at the UK Infrared Telescope yielded a K ($2.2\ \mu\text{m}$) magnitude of 16.8, corresponding to a luminosity of $1.6 \times 10^{12} L_{\odot}$ at $6,695\ \text{\AA}$. Thus 10214+4724 lies 1.5 magnitudes above the K - z relation for radiogalaxies. But the optical and ultraviolet radiation account for $<1\%$ of the total bolometric output. The object was observed by R. Cutri and P. Eisenhardt in a 17-minute integration at $10\ \mu\text{m}$ on 27 January 1991 at the Multiple Mirror Telescope, but not detected: a 3σ upper limit of 12 mJy was obtained (P. Eisenhardt, personal communication). The overall energy distribution from $3,000\ \text{\AA}$ to radio wavelengths is similar to that of 09104+4109 (ref. 7).

The continuum spectrum from $1,400$ to $3,000\ \text{\AA}$ is similar to those of the high-redshift radiogalaxies, though rather redder than most of these. It is roughly of the form $f_{\nu} \propto \nu^{-1.7}$. Such a spectrum can be modelled as a continuum spectrum from a quasistellar object, scattered by gas or dust, or as a population of luminous young stars suitably reddened. If dust has considerably extinguished or scattered the spectrum, it must be relatively free of the component responsible for the feature at $2,200\ \text{\AA}$ in the local interstellar medium.

As no narrow emission-line galaxies with $z > 0.8$ have been found in optical redshift surveys of more than 300 galaxies to limits two magnitudes fainter than 10214+4724 (refs 8–10), they must constitute less than 1% of the faint galaxy population. On the other hand, we have surveyed only 1/60 of the sky to 0.2 Jy at $60\ \mu\text{m}$, so many other examples may remain to be found. From our Faint Source galaxy survey to date, we estimate the surface density of high-redshift ($z > 1$) galaxies brighter than 0.2 Jy at $60\ \mu\text{m}$ and with $R \leq 20.5$ to be ≤ 0.01 per square degree.

TABLE 1 Infrared, optical and radio properties of IRAS 10214+4724

	Right ascension	Declination
IRAS FSS position	10 h 21 min 29.9 s	+47°24'42" (1950.0)
IRAS Addscan position	10 h 21 min 29.9 s	+47°24'20"
Object F	10 h 21 min 31.12 s	+47°24'23.9" ($\pm 0.5''$)
VLA source	10 h 21 min 31.14 s	+47°24'22.9" ($\pm 0.3''$)

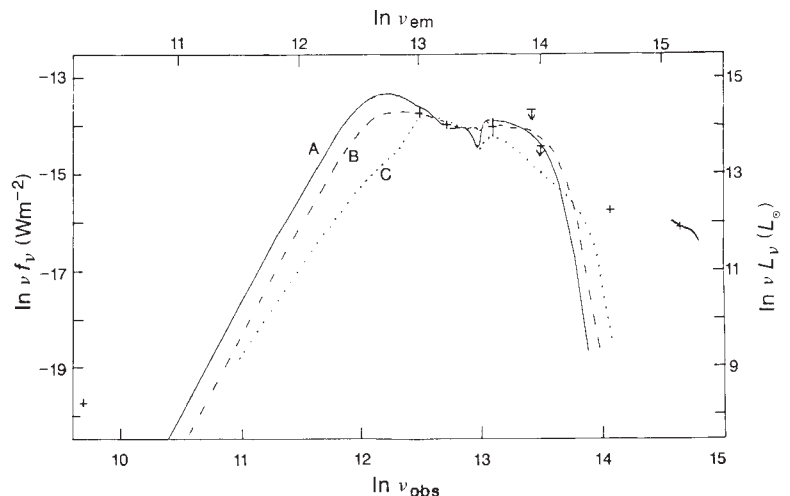
IRAS fluxes from Addscans: $S_{100} = 550 \pm 150$ mJy, $S_{80} = 190 \pm 40$ mJy, $S_{25} = 75 \pm 31$ mJy, $S_{12} < 90$ mJy. Radio flux: $S(4.86\ \text{GHz}) = 0.36$ mJy. K-band ($2.2\ \mu\text{m}$): $K = 16.8$. Optical magnitude: Gunn $r = 20.9$. Assuming $(R-V) = 0$, this implies Johnson $R = 20.5$.

If 10214+4724 represents a luminous primaeval phase common in massive galaxies at $z = 2-3$, the absence of similar examples in deeper optical surveys is surprising considering that it would take at least 10^6 years to generate the obscuring dust¹¹.

Whether or not the emission lines are interpreted as evidence for an active nucleus, this tells us little about what is powering the far infrared emission. The spectral index ($\alpha = -d(\log f_{\nu})/d(\log \nu)$) between the two clear detections at rest wavelengths of 31 and $19\ \mu\text{m}$ is $\alpha = 2.1 \pm 0.7$. Taking into account the marginal detection at $25\ \mu\text{m}$ and the upper limit at $10\ \mu\text{m}$, the spectral index between 3 and $31\ \mu\text{m}$ is ≥ 1.7 , steeper than most quasars in this region, for which $\alpha \approx 1.0$ (refs 12, 13). The steep drop in flux between the observed wavelengths of $100\ \mu\text{m}$ and $6\ \text{cm}$, similar to that seen in radio-quiet quasars and active galactic nuclei^{12,14}, is suggestive of a thermal origin, though a self-absorbed synchrotron spectrum cannot be ruled out. As argued by Kleinmann *et al.*⁷ in the case of 09104+4109, however, a nonthermal model is implausible because of the very compact size implied ($\leq 10^{-5}$ arcsec in the case of our object). Observations at radio wavelengths will provide an important test of this hypothesis.

The observed colour ratios of fluxes at 25, 60 and $100\ \mu\text{m}$ for 10214+4724 lie outside the range found for IRAS galaxies with high-quality, unresolved fluxes¹⁵. But when the effects of the redshift are allowed for, the infrared energy distribution is similar to that of nearby starburst galaxies. Recently A. Efstathiou and M.R.-R. (manuscript in preparation) have constructed an improved starburst galaxy model based on a new multiple-grain model¹⁶. To fit the characteristic spectrum of starburst galaxies at $12-100\ \mu\text{m}$, and the detailed spectrum available for the starburst component in NGC1068, models with a uniform density distribution were preferred. The model adopted for normal IRAS starburst galaxies is a good fit to the rest-frame spectrum of 10214+4724 (curve A in Fig. 3). The dust mass involved, however, is $1.0 \times 10^{10} M_{\odot}$, far in excess of the dust

FIG. 3 Radio-to-ultraviolet spectrum of 10214+4724, with the far infrared emission compared with redshifted theoretical models: (model A) solid curve: standard Starburst model (A. Efstathiou and M.R.-R., manuscript in preparation) with $T_s = 40,000\ \text{K}$, $\tau_{\text{uv}} = 200$, $r_1/r_2 = 0.002$; (B) broken curve: as (A) but $\tau_{\text{uv}} = 100$, $r_1/r_2 = 0.003$; (C) dotted curve: Seyfert model of Rowan-Robinson and Crawford¹⁵ with power-law illuminating source, $\alpha = 0.7$, $\tau_{\text{uv}} = 75$, $T_1 = 500\ \text{K}$, $r_1/r_2 = 0.0215$. The bolometric luminosities of the three models are 7.2×10^{14} (A), 4.7×10^{14} (B) and 1.9×10^{14} (C) $L_{\odot}$ (for $H_0 = 50$). The left-hand and lower scales give the observed fluxes and frequencies ($\log_{10}(\nu f_{\nu})$ versus $\log_{10} \nu_{\text{obs}}$ (Hz)) and the right-hand and upper scales give the emitted quantities ($\log_{10}(\nu L_{\nu})$ versus $\log_{10} \nu_{\text{em}}$).



mass found in a massive spiral like our own ($2 \times 10^8 M_\odot$), and five times greater than the entire mass of heavy elements in the disk of our Galaxy ($2 \times 10^9 M_\odot$, assuming a total disk mass of $10^{11} M_\odot$ and a heavy element abundance of 2%). Such a model could only survive if 10214+4724 represented a primaeval phase of an exceptionally massive ($\geq 10^{12} M_\odot$) galaxy.

A less extreme model is shown as curve B in Fig. 3, with a dust mass of $1.5 \times 10^9 M_\odot$. This could correspond to the process of formation of the Population I heavy elements in a massive spiral galaxy. The infrared power could be generated by 10^9 O5 stars ($M = 40 M_\odot$), with a total mass of $4 \times 10^{10} M_\odot$. A characteristic heavy element yield of 30% would then be consistent with the required dust mass. This model is marginally inconsistent with the 10- μm upper limit, although the fit could be improved by reducing the maximum grain temperature T_1 (assumed to be 1,000 K in models A and B). This would not change the estimate of dust mass. For a starburst origin of the far-infrared emission, the galaxy would have to be older than 10^6 years for considerable amounts of heavy elements to have been dispersed from stars, and could not be much older than 3×10^7 years to avoid an absurdly high energy budget. Because of the high optical depth that has to be assumed for the dust cloud, the emergent spectrum is not strongly dependent on the nature of the illuminating source. Thus from the infrared observations alone, it is equally possible that we are seeing a shrouded quasar of high luminosity. For a uniform density distribution, however, the estimates of dust mass are unchanged, and they are so excessive that such models are inconsistent with the idea of a quasi-stellar object switching on in a pre-existing normal galaxy.

We have also investigated whether models of the type used to fit the mid-infrared spectra of Seyfert galaxies¹⁵ could be consistent with the spectrum of 10214+4724. These models involve a $\rho(r) \propto r^{-1}$ distribution of dust in the narrow-line region of the active galactic nucleus. Model C in Fig. 3 shows a fit with the model used for the core of NGC1068, for Mk231, NGC1275 and other Seyfert galaxies¹⁵. The model has optical depth at $1000 \text{ \AA}$ $\tau_{\text{uv}} = 75$ and $T_1 = 500 \text{ K}$. This is an acceptable fit to the observations. The dust mass is $4 \times 10^8 M_\odot$, somewhat higher than has been seen in any galaxy to date. The main support for this idea that we are seeing an embedded quasi-stellar object switch-on in a pre-existing galaxy lies in the very high total luminosity and in the high excitation of the narrow emission lines. Such a model was proposed for 09104+4109 by Kleinmann *et al.*⁷, who argued that the optical and near-infrared radiation represents a relatively unobscured leak of the underlying nonthermal continuum through a dust cloud with a covering factor of 98%. A difficulty with the hypothesis is the weakness of Lyman- α .

The other hypothesis that provides an acceptable fit to the infrared spectrum is an embedded starburst associated with the formation of the Population I heavy elements in a massive spiral galaxy. Submillimetre observations will provide a crucial test of these hypotheses. Whatever the source of illumination, this may be the first evidence for a large mass of dust in a galaxy at $z > 2$, implying that considerable chemical enrichment has already occurred in the galaxy at this early epoch. It will be highly desirable to obtain more detailed observations of this unique object and to search for other objects of this type.

Note added in proof. The total luminosity of IRAS 10214+4724 exceeds that of the previously most luminous known object in the universe, the quasar S5 0014+81 (refs 17–19), for which the 0.9–2.5 μm luminosity is $2.2 \times 10^{14} L_\odot$ ($H_0 = 50$, $\Omega_0 = 1$). □

6. Binette, L., Dopita, M. A., D'Odorico, S. & Benvenuti, P. *Astr. Astrophys.* **115**, 315–320 (1982).
7. Kleinmann, S. G. *et al. Astrophys. J.* **328**, 161–169 (1988).
8. Broadhurst, T. J., Ellis, R. S. & Shanks, T. *Mon. Not. R. astr. Soc.* **235**, 827–856 (1988).
9. Colless, M. N., Ellis, R. S., Taylor, K. & Hook, R. N. *Mon. Not. R. astr. Soc.* **244**, 408–423 (1990).
10. Lilly, S. J., Cowie, L. L. & Gardner, J. P. *Astrophys. J. Suppl. Ser.* (in the press).
11. Cowie, L. L. *The Post-Recombination Universe* (eds Kaiser, N. & Lazenby, A. J.) 1–18 (Kluwer, Dordrecht, 1988).
12. Sanders, D. B. *et al. Astrophys. J.* (in the press).
13. Elvis, M. *et al. Astrophys. J. Suppl. Ser.* (submitted).
14. Lawrence, A. *et al. Mon. Not. R. astr. Soc.* **248**, 91–100 (1991).
15. Rowan-Robinson, M. & Crawford, J. *Mon. Not. R. astr. Soc.* **238**, 523–538 (1991).
16. Rowan-Robinson, M. *Interstellar Medium in Galaxies* (eds Thronson, H. A. Jr & Shull, J. M.) 121–150 (Kluwer, Dordrecht, 1990).
17. Kuhr, H. *et al. Astrophys. J.* **275**, L33–L37 (1983).
18. Kuhr, H. *et al. Astrophys. J.* **284**, L5–L8 (1984).
19. Steidel, C. C. & Sargent, W. L. W. *Astrophys. J.* **313**, 171–184 (1987).

ACKNOWLEDGEMENTS. We thank P. McCarthy, R. Cutri, P. Eisenhardt, C. Benn and the staff at UKIRT and the WHT for observations of 10214+4724. T.J.B., A.L., R.G.M., S.J.O., A.N.T. and W.S. acknowledge support from SERC. Research at IPAC is carried out under a contract with NASA.

Optical transitions in diamond at ultrahigh pressures

H. K. Mao & R. J. Hemley

Geophysical Laboratory and Center for High-Pressure Research,
Carnegie Institution of Washington, 5251 Broad Branch Road NW,
Washington DC 20015, USA

MANY technological and scientific applications of diamond arise from its unique properties, which include optical transparency in the ultraviolet to infrared, electrical insulation, thermodynamic stability, and unsurpassed strength and hardness¹. Here we describe optical studies on diamond at ultrahigh pressures (to above 300 GPa) which show that these properties are affected by such stresses. Our spectroscopic measurements show that the optical absorption edge shifts from ultraviolet to red with increasing pressures, and Raman scattering measurements show evidence for new structural transitions, associated with large macroscopic deformation, beginning at a pressure of approximately 150 GPa. The changes are reversible and are associated with intense luminescence peaks at 2.0–2.2 eV under 458–514-nm radiation. Our results may be related to the onset of band-gap closure in the approach to a new high-pressure phase. These spectral features must also be taken into account when diamond is used as optical windows for ultrahigh-pressure investigations^{2–5}.

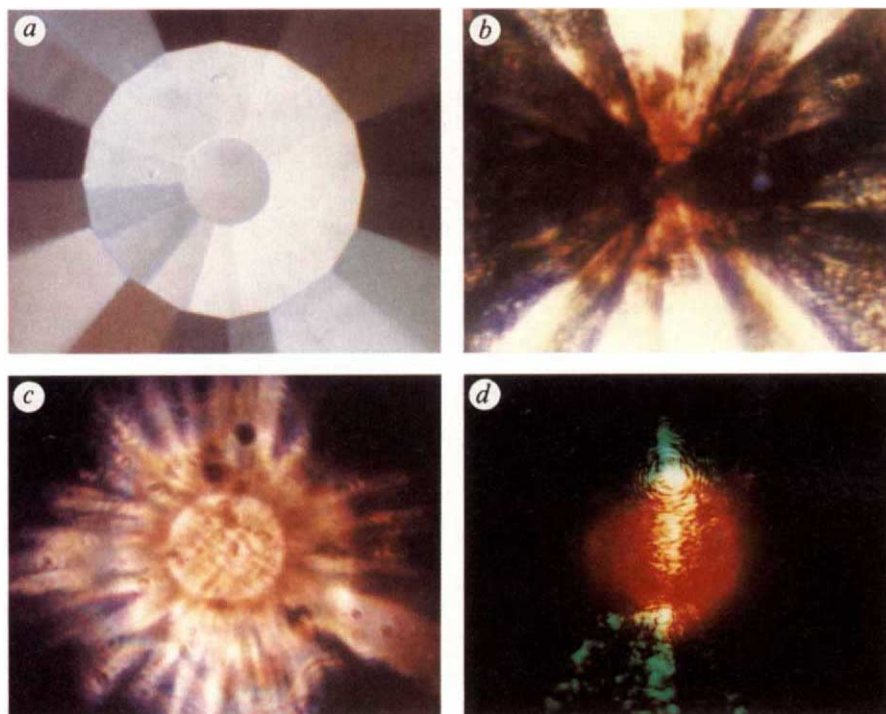
Electronic-structure calculations predict that diamond is stable to at least 1,200 GPa (12 Mbar) hydrostatic pressure^{6–8} but may transform under considerably lower stresses (400 GPa) if the load is uniaxial⁹. It remains an open question whether such transformations set an upper limit on the pressures that can be reached with diamond cells. Plastic flow of diamond at room temperature was first demonstrated¹⁰ in 1978 at sample pressures of 172 GPa. Simultaneously, the highly stressed anvil tip developed a light yellowish brown coloration. One important feature of the diamond cell is the transparency of the diamond windows to large regions of the electromagnetic spectrum. If diamond develops new optical features associated with band-gap closure⁹, spectroscopic characterization of materials under static ultrahigh pressure conditions will be affected^{11,12}. The aggregated nitrogen in type Ia diamonds used for ultrahigh pressure studies¹⁰ gives rise to a variety of absorption and luminescence bands in the visible and ultraviolet under low stresses¹. It is important to examine changes in these impurity levels at very high pressures. Changes in bonding properties and phase transformations can also be examined by vibrational Raman spectroscopy, which has been a definitive probe for carbon phases at lower pressures^{13–17}.

We used selected type Ia diamonds, which are transparent in the visible, are free of sharp fluorescence peaks and have broad-band fluorescence weaker than the second-order Raman bands¹¹. In one experiment, type IIa diamonds were used. Here we

Received 12 February; accepted 14 May 1991.

1. Rowan-Robinson, M. *Advances in Space Research* **11**, 247–254 (1991).
2. Chambers, K. C. & McCarthy, P. J. *Astrophys. J.* **354**, L9–L12 (1990).
3. Spinrad, H. *et al. Astrophys. J.* **299**, L7–L11 (1985).
4. Spinrad, H. *Publ. astr. Soc. Pacific* **98**, 269–284 (1985).
5. Blair, W. P., Raymond, J. C., Danziger, J. & Matteucci, F. *Astrophys. J.* **338**, 812–823 (1989).

FIG. 1. Micrographs of bevelled diamond anvils⁴ at ultrahigh pressures viewed through the anvils at the highly stressed tip. A white-light quartz halogen lamp was used for all except the last micrograph. *a*, Diamond anvil at zero pressure (stress-free) viewed in transmitted light. Cross-polarization was used to show the bevelling and faceting of the anvil. *b*, Diamond anvil at high stress viewed in reflected light (295 K). The sample consisted of an opaque mixture of platinum and caesium iodide at a peak pressure of 302 GPa (as determined by X-ray diffraction from the platinum¹⁸). The anvil-tip diameters of the upper and lower diamonds were 50 and 25 μm . *c*, View through a pair of anvils in transmitted light at a pressure of ~ 300 GPa (77 K). The sample was neon, which remains transparent under these pressures. The larger flat has a diameter of 100 μm . The small flat of 50 μm and the bevel are not observed because of the elastic deformation. A cross-hatched texture seen at the tip indicates plastic deformation. *d*, Diamond at ~ 300 GPa (77 K) under 488-nm laser excitation showing the intense red luminescence from within the anvil near the highly stressed tip. The green colour is the scattered laser light.



summarize the results from 30 different experiments carried out above 150 GPa. To distinguish unequivocally between the effects of pressure-induced optical changes in the diamond windows and in the samples, we used fourteen different samples: Ne, NaCl, ruby, SiO_2 , H_2 , D_2 , N_2 , CsI, Fe, Fe-Ni alloy, FeO, W, Pt and Re, representing different classes of materials with a wide range of properties. Solid neon remains a large band-gap insulator and is transparent over the entire range of this study⁴; it is thus ideal for studying the transmission spectra of diamond anvils. Iron, iron-nickel, tungsten, platinum and rhenium are opaque metals over the entire spectral and pressure ranges of the present study, and thus have high reflectivity. Other samples are expected to transform from insulator or semiconductor to metal at pressures below 300 GPa (refs 4 and 5). In particular, CsI, which is a transparent insulator at zero pressure, becomes an opaque metal above 130 GPa (refs 5 and 18). Each of the optical phenomena reported here for diamond has been confirmed with at least three or four types of samples.

In these experiments the sample serves as a thin film (typically $\sim 1\text{--}2\ \mu\text{m}$) of a softer material which prevents brittle fracture of the diamonds. Examples of changes for both optically opaque and transparent samples are shown in Fig. 1. With increasing pressure, the anvil tips change from transparent (Fig. 1*a*) to yellow, to brown, eventually becoming deep red. In Fig. 1*b*, the red colour is apparent in reflected light in the region of peak pressure, which was determined¹⁸ by *in situ* X-ray diffraction of the opaque sample (CsI) to be 302 GPa at 295 K. Further information is available from the micrograph of solid neon under a similar load (pressure ≥ 250 GPa) and at 77 K (Fig. 1*c*). The diamond tip is bright yellow. A distinct lamination or striated pattern is apparent at the diamond surface. This pattern indicates plastic deformation of the anvil tip¹⁰, which has been documented for naturally deformed diamonds¹⁹, but not previously identified *in situ* at high pressure. In other experiments, such macroscopic deformation associated with the coloration was not observed.

To characterize the colour changes further, we studied the variation of the absorption and luminescence spectra from different regions of the anvils. We measured absorption from direct transmission using an aperture of 2- μm diameter. For laser-excitation studies, we used a $\sim 135^\circ$ scattering configura-

tion¹² with a 2- μm -diameter sampling aperture and beams from both Ar^+ and Kr^+ ion lasers focused to comparable diameter (to $\sim 2\ \mu\text{m}$). The off-axis excitation and spatial filtering technique with the sampling aperture allows us to produce a three-dimensional map of the optical characteristics (luminescence and Raman spectrum) of the highly strained diamond. The measurements showed that the red colour is caused not only by increased absorption in the blue but also by intense luminescence in the red (Fig. 2*a*). There is an appreciable increase in the broad visible luminescence when the sample-diamond interface is probed, beginning at a sample pressure of 100–150 GPa^{12,20}. At ~ 300 GPa, the intensity increases markedly (by up to a factor of 10^3), and a sharp red peak at $\sim 2\ \text{eV}$ (620 nm) appears^{21,22}. The efficiency of the fluorescence is extremely high, as is evident in the unfiltered micrograph of Fig. 1*d*, which shows the intensity of the bright red luminescence of the diamond anvil to be comparable to that of the scattered and reflected part of the blue-green excitation-laser beam at 488 nm. The absorption measurements show that there is a considerable red shift in the absorption edge at 100 GPa; at pressures of about 300 GPa, absorption throughout the visible is observed. We find, however, that the shift in the effective absorption edge of the diamond depends on the stress on the anvil: a larger shift is documented for the weaker samples (Fig. 2*b*). The intensity of the broad-band luminescence is directly correlated to the absorption in the blue, and is much weaker ($\times 10^{-3}$) when the excitation wavelength changes from 488.0 to 568.2 nm at 200 GPa.

Spatially resolved micro-Raman spectra at the anvil tips provide direct information about the effect of high stress on the diamonds. At zero stress, the triply degenerate (zone-centre) T_{2g} optical phonon of the diamond at $1,333\ \text{cm}^{-1}$ is clearly visible (Fig. 3*a*). Previous measurements^{13–16} of the stress dependence of this band (in both diamond samples and diamond anvils) have been limited to <40 GPa. In Fig. 3*b*, we show the spectrum of this band measured at the centre of the flat of a sample of solid nitrogen at 160 GPa, where the frequency is $1,600\ \text{cm}^{-1}$. A well defined peak is observed; considerable broadening of the band is observed when the measurement is carried out off centre, as a result of the splitting of the triply degenerate T_{2g} peak in the triaxial stress field of the diamond. At sample

pressures of 150–200 GPa, a new Raman peak appears at 590 cm^{-1} (Fig. 3c). The new band seems to be associated with the appearance of yellow fluorescence, and occurs abruptly with increasing load, but the onset pressure and intensity of the peak appears to depend on the individual diamond and stress conditions of the anvil. The feature does not correlate with known Raman bands in previously characterized carbon phases, including lonsdaleite (hexagonal diamond), amorphous or glassy carbon¹⁷. Notably, careful measurements at the sample–anvil interface indicate that the intensity of the band can overwhelm the diamond T_{2g} mode. The bands were reversible on releasing the stress, but there is a large degree of hysteresis: the peak can be preserved to pressures as low as 50 GPa. The tips of the diamonds also show stronger luminescence after loading.

The abrupt change in both the luminescence and Raman spectra indicates that considerable changes in bonding properties and structure of diamond occur under these conditions. The luminescence may be due to pressure-induced electronic changes in deep-level impurity centres. If so, the new Raman band may be associated with localized vibrational modes at these centres^{1,23}. Brown coloration and luminescence at $\sim 2\text{ eV}$ documented in naturally deformed diamonds have been attributed to dislocations arising from deformation¹⁹. Although these optical features are similar to the present *in situ* observations, the changes observed here are reversible and represent a different type of transformation. We therefore consider that the new bands may be caused by changes in the intrinsic electronic structure of diamond. This interpretation is supported by our observations that the stress dependences of the absorption edge

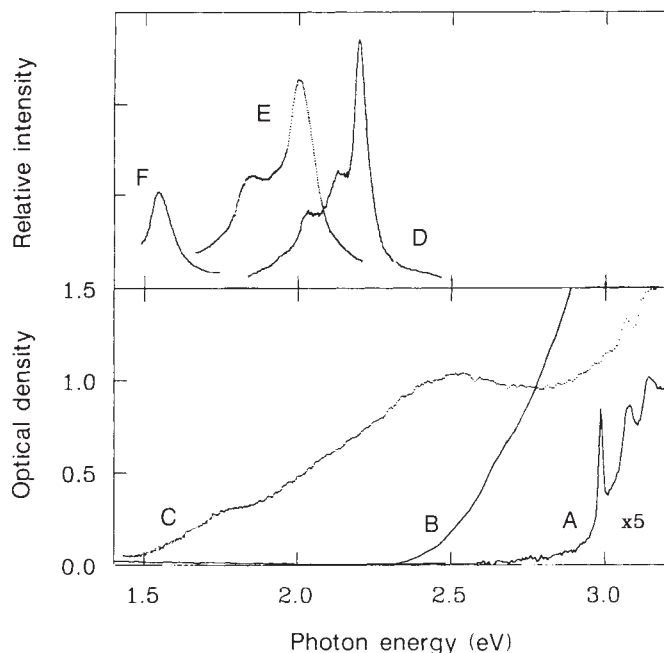


FIG. 2 Effect of pressure on absorption (lower panel) and fluorescence (upper panel) spectra from diamond anvils. Lower panel, A, absorption spectrum through a pair of type Ia diamonds at zero pressure (295 K) showing the absorption by nitrogen beginning near 3 eV; B, absorption edge of diamond obtained with a weak, or quasihydrostatic, sample (hydrogen²⁶) at a pressure of 161 GPa (295 K); C, diamond at ~ 300 GPa with a stronger sample (NaCl²¹ at 295 K). In both samples the thickness was estimated to be $\leq 1\text{ }\mu\text{m}$. Upper panel, D, fluorescence spectrum from the diamond tip at sample pressures of 300 GPa with 458.0-nm excitation (see Fig. 1d; 77 K); E, same diamond and stress conditions with 514.5-nm excitation; F, near-infrared peak excited with 647.1-nm radiation (77 K); a similar peak also appears with blue and green excitation. Experiments carried out with different samples indicate that the absorption and fluorescence features shown here are intrinsic to the stressed diamond anvil and do not arise from the sample. Peak pressures were estimated from the pressure profile determined from ruby fluorescence or from X-ray diffraction measurements on metal pressure standards⁴ (or both).

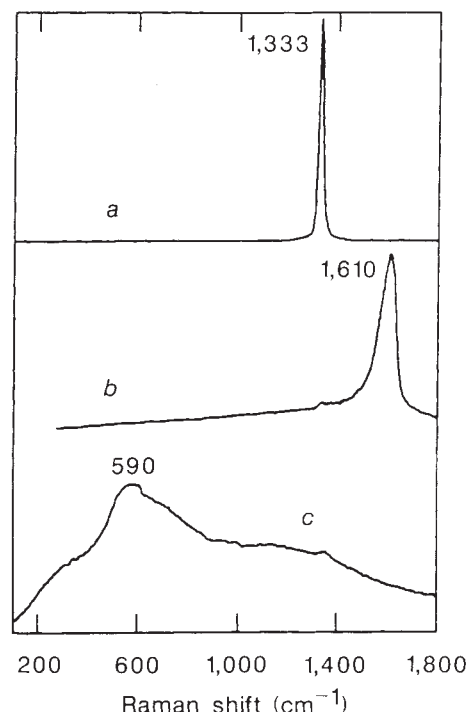


FIG. 3 Raman spectrum of diamond at the sample–anvil interface. a, Zero pressure. b, 160 GPa (solid N_2 sample, 295 K). Broad low-frequency Raman bands of nitrogen are not apparent on this scale. c, ~ 300 GPa (solid neon sample, 77 K). The 590 cm^{-1} band was observed in nine experiments with H_2 , D_2 , Ne, CsI, NaCl and SiO_2 samples above 150 GPa. It was not observed in six other experiments that terminated at 150–200 GPa. The band shows resonance enhancement with blue-green excitation, and is independent of temperature from 77 to 295 K. Weaker Raman bands at ~ 900 and $2,000\text{ cm}^{-1}$ have also been observed. In addition, fluorescence bands appear in this region, but they can be distinguished from Raman peaks by varying the laser excitation wavelength.

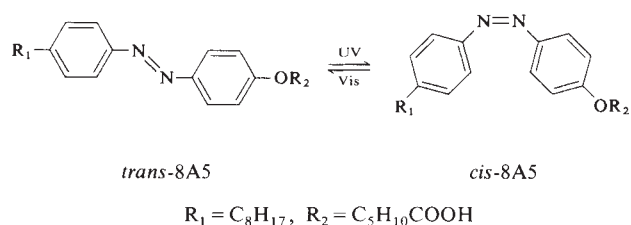
in type I and type II diamonds, which have very different impurities, are similar above 100 GPa with samples of similar strength. A considerable tetragonal distortion of the diamond structure is expected at the anvil tip, which is compressed along [100] in our experiments. For a principal normal stress of 300 GPa, for example, the ratio of cell dimensions c/a is expected to be ~ 0.8 . Under such conditions, an increase is expected in the number of Raman and electronic transitions allowed by symmetry. We also point out that there is a close similarity between the new Raman feature and the one-phonon density of states of diamond (which has a broad peak centred at 600 cm^{-1})²⁴. Such a correlation could imply a breakdown in crystalline selection rules, resulting perhaps from growth of defects (stacking faults) at the anvil tips associated with macroscopic flow of the diamond, or from the formation of a new high-density amorphous form of carbon.

There has been considerable interest in the possibility of producing high-density polymorphs of diamond, including metallic forms, by applying high hydrostatic pressure^{6–9}. Density-functional theory calculations for diamond under large uniaxial strains⁹ predict that the band gap closes at stresses of 400–450 GPa. In the metallic state calculated for large [110] and [111] strains, the T_{2g} phonon is predicted to become unstable⁹. The loss of intensity of this band in our spectra may be related to this instability. The observation of new Raman bands indicates that there may be new insulating (or semiconducting) phases of carbon which are stable at very high hydrostatic or nonhydrostatic stresses but which have not yet been considered in theoretical calculations. The red shift of the absorption edge is consistent with a gradual closing of the band gap between valence and conduction bands at much higher

pressure; similar results have been obtained by other workers (Y. K. Vohra and A. L. Ruoff, personal communication). But accurate prediction from these data of the transition pressure to a metallic state of diamond is not possible in view of the wide range of the extrapolation and the evidence for abrupt changes in the optical properties and structure documented here.

Although diamond anvils start to show signs of structural transition or plastic deformation at pressures as low as 150 GPa, the anvils remain intact, and can be compressed consistently to still higher pressures (>250 GPa (ref. 4)). Therefore, these changes, regardless of whether they are intrinsic to diamond or defect-related, do not set an absolute limit on the useful working range of the diamond cell. Furthermore, optical properties of samples under these conditions (such as the Raman and optical spectra of solid hydrogen^{25,26}) can still be studied through the coloured diamond windows, as long as the spectrum of the window is well characterized. □

There have been several studies of the regulation of membrane properties by photochemical reactions of azobenzene derivatives, by incorporating them into Langmuir-Blodgett films¹, vesicles², microcapsules³ and liquid-crystal⁴ and polymeric⁵ films. The derivative that we have used here is 4-octyl-4'-(5-carboxy-pentamethylene-oxy) azobenzene (8A5), which undergoes reversible *cis-trans* isomerization in solution on irradiation with ultraviolet and visible light:



The experimental set up is schematically shown in Fig. 1. The resistance of the modified black lipid membrane (BLM), as estimated from a straight-line fit to the current-voltage relation in the cell, was $(4-40) \times 10^9 \Omega$. The resistivity of the modified BLM (surface area $\sim 1.8 \text{ mm}^2$) was $(0.7-7.2) \times 10^8 \Omega \text{ cm}^2$, which is close to that for the unmodified membrane from egg lecithin⁶. At an applied potential, V_{app} , of 10 mV, the modified BLM was illuminated with visible light (wavelength 450 nm) for ~ 30 min, followed by irradiation with ultraviolet light (365 nm) for ~ 10 min. When irradiation was then changed back to visible light, a sharp current signal in the negative direction, with a decay time of ~ 50 s, was observed. Subsequent irradiation with

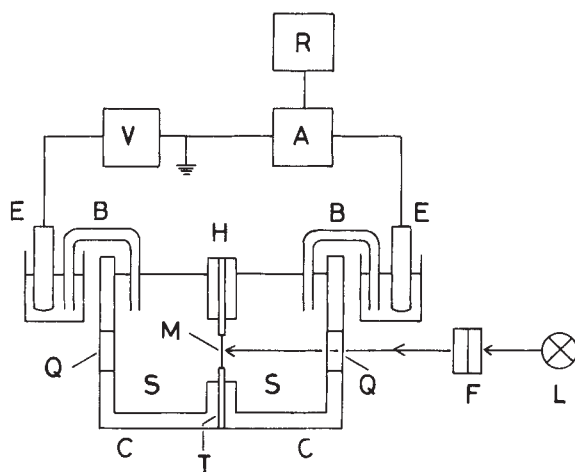


FIG. 1 Schematic representation of the experimental set-up. The BLM cell consisted of two Teflon compartments (C) having quartz windows (Q) for illumination and inspection. In the Teflon septum (T) (thickness 0.1 mm) which was fixed by the septum holder (H) and separated the two compartments, there was a hole of diameter 1.5–2 mm, in which the black lipid membrane was formed. The compartments were filled with a 0.01 M aqueous solution of KCl (S) and connected with two Ag/AgCl electrodes (E) by KCl salt bridges (B). The membrane was fabricated by a brush technique⁸⁻¹⁰. The membrane-forming solution consisted of 50 mg of egg lecithin (Sigma chemicals) for the unmodified BLM, or of 50 mg of egg lecithin plus 4.5–4.6 mg of 8A5 (Dojin Kagaku) for the modified BLM, in 0.65 ml of *n*-decane:CHCl₃ solution (10:3 by volume). The voltage, $V_{\text{app}} = 0-50$ mV, from a d.c. power supply (V) was applied to the cell, and the current was measured by a picoammeter (A) and monitored on the recorder (R). To illuminate the BLM, we used a 500-W high-pressure mercury lamp (L) in combination with suitable glass filters (F). The incident photon number to the cell at $\lambda = 365$ nm and 450 nm, as determined by ferrioxalate actinometry¹¹ and a photometer, were 2.4×10^{15} and $9.6 \times 10^{15} \text{ cm}^{-2} \text{ s}^{-1}$, respectively. The Ag/AgCl electrodes were carefully screened to cut out incident light. The freshly prepared BLM was preconditioned by irradiation with 450-nm light at $V_{\text{app}} = 10$ mV.

Received 25 March; accepted 2 May 1991.

1. Davies, G. *Diamond* (Hilger, Bristol, 1984).
2. Jayaraman, A. *Rev. mod. Phys.* **55**, 65–108 (1983).
3. Seal, M. *High Temp.-High Pressures* **16**, 573–579 (1984).
4. Mao, H. K. in *Simple Molecular Systems at Very High Densities* (eds Polian, A., Loubeyre, P. & Bocarra, N.) 221–236 (Plenum, New York, 1989).
5. Jeanloz, R. *Ann. Rev. phys. Chem.* **40**, 237–259 (1989).
6. Yin, M. T. & Cohen, M. *Phys. Rev. Lett.* **50**, 2006–2009 (1983).
7. Yin, M. T. *Phys. Rev.* **B30**, 1773–1776 (1984).
8. Biswas, R. Martin, R. M., Needs, R. J. & Nielsen, O. H. *Phys. Rev.* **B30**, 3210–3213 (1984).
9. Nielsen, O. H. *Phys. Rev.* **B34**, 5808–5819 (1986).
10. Mao, H. K., Bell, P. M., Dunn, K. J., Chrenko, R. M. & Devries, R. C. *J. appl. Phys.* **50**, 1002–1009 (1979).
11. Adams D. M. & Sharma, S. K. *J. Phys. E: Sci. Instrum.* **10**, 680–682 (1977).
12. Hemley, R. J., Bell, P. M. & Mao, H. K. *Science* **237**, 605–612 (1987).
13. Sharma, S. K., Mao, H. K., Bell, P. M. & Xu, J. A. *J. Raman Spectrosc.* **16**, 350–352 (1985).
14. Hanfland, M., Syassen, K., Fahy, S., Louie, S. G. & Cohen, M. L. *Phys. Rev.* **B31**, 6896–6899 (1985).
15. Boppert, H., van Straaten, J. & Silvera, I. F. *Phys. Rev.* **B32**, 1423–1425 (1985).
16. Goncharov, A. F., Makarenko, I. F. & Stishov, S. M. *JETP Lett.* **41**, 184–187 (1985).
17. Knight, D. S. & White, W. B. *J. Mater. Res.* **4**, 385–393 (1989).
18. Mao, H. K., Wu, Y., Hemley, R. J., Chen, L. C., Shu, J. & Finger, L. W. *Science* **246**, 649–651 (1989).
19. Orlov, Y. L. *The Mineralogy of Diamond* (Wiley, New York, 1977).
20. Goettal, K. A., Mao, H. K. & Bell, P. M. *Rev. Sci. Instrum.* **56**, 1420–1427 (1985).
21. Hemley, R. J. & Mao, H. K. *Ann. Report Director Geophys. Lab. 1988–1989*, 105–108 (1989).
22. Vohra, Y., Xia, H., Luo, H. & Ruoff, A. F. *Appl. Phys. Lett.* **57**, 1007–1009 (1990).
23. Davies, G. & Hamer, M. F. *Proc. R. Soc. Lond. A* **348**, 285–298 (1976).
24. Dolling, G. & Cowley, R. A. *Proc. Phys. Soc. (Lond.)* **88**, 463–494 (1966).
25. Hemley, R. J. & Mao, H. K. *Phys. Rev. Lett.* **61**, 857–860 (1988).
26. Mao, H. K. & Hemley, R. J. *Science* **244**, 1462–1465 (1989).

ACKNOWLEDGEMENTS. We thank J. F. Shu and M. Hanfland for help with the experiments, and C. Meade and F. R. Boyd for very useful discussions. This work was supported by the NSF, NASA and CIW.

Photoelectric response of a black lipid membrane containing an amphiphilic azobenzene derivative

H. Fujiwara & Y. Yonezawa*

Department of Industrial Chemistry, Faculty of Engineering, Kyoto University, Yoshida, Kyoto 606, Japan

LIPID membranes containing dye molecules that respond to light, for example by photoisomerization, may provide the basis for artificial visual systems for use in optoelectronic devices and optical transducers. Here we describe such a membrane/dye system that generates an electrical signal in response to irradiation within a specific range of wavelengths. We incorporated an amphiphilic azobenzene derivative into the planar black lipid membrane formed from egg lecithin. When a d.c. voltage is applied across the modified membrane, transient current pulses are induced by alternate irradiation with ultraviolet and visible light. We interpret this photoelectric response as the result of changes in the membrane capacitance caused by *cis-trans* photoisomerization of the azobenzene derivative.

* To whom correspondence should be addressed.

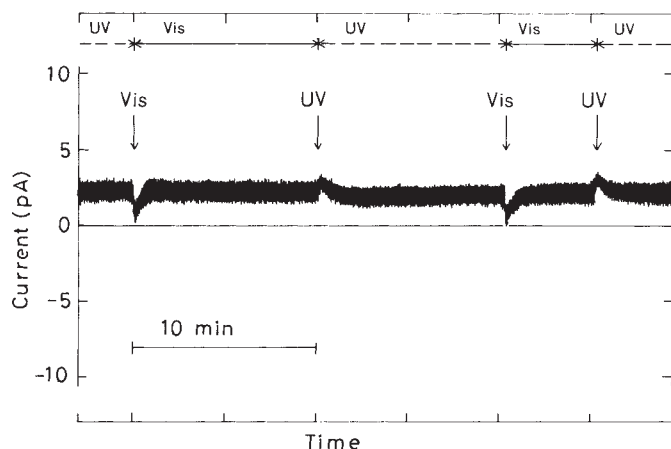


FIG. 2 Time dependence of the change in direct current on continuous illumination of the modified BLM with ultraviolet ($\lambda = 365$ nm) and visible ($\lambda = 450$ nm) light. $V_{app} = 10$ mV. Duration of ultraviolet (---) and visible (—) light is shown at the top of the figure. The arrows indicate the change from ultraviolet light to visible or vice versa.

ultraviolet light produced a current signal in the positive direction (decay time ~ 100 s). A typical result is given in Fig. 2. The unmodified BLM did not show any photoresponse. The peak photocurrent on changing from ultraviolet to visible irradiation (wavelength $\lambda = 400$ – 520 nm) was measured and normalized to the current produced at uniform light intensity (the response to this change is larger than the response to the reverse change from visible ultraviolet). The action spectrum (normalized

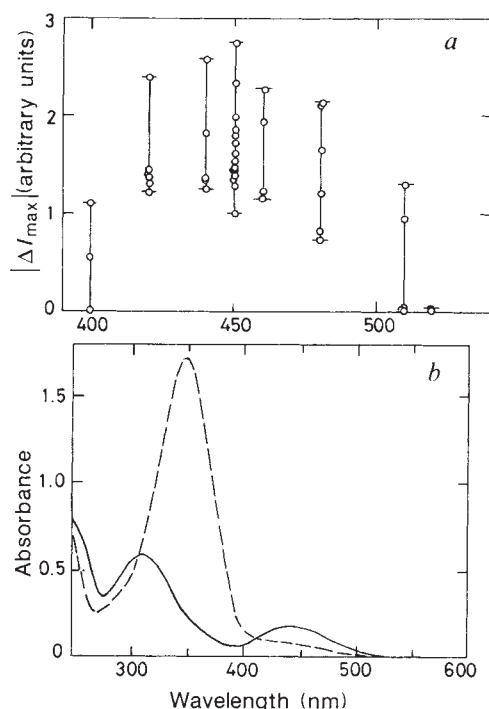


FIG. 3 Action spectrum of photocurrent in the negative direction (A) and absorption spectra of 0.1 mM chloroform solution of 8A5 (B). *a*, Action spectrum $|\Delta I_{max}|$ against wavelength, in which $|\Delta I_{max}|$ denotes the magnitude of peak photocurrent normalized to uniform light intensity. $V_{app} = 10$ mV. *b*, Absorption spectra of 8A5 solution. Dashed line, *trans*-8A5, after irradiation with 450 nm light; solid line, *cis*-8A5, after irradiation with 365 nm light. Light-path length = 10 mm.

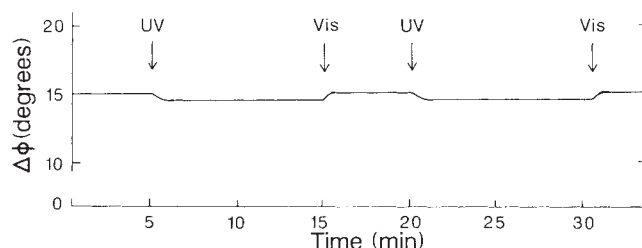


FIG. 4 Time dependence of the phase difference, $\Delta\phi$, in the a.c. current on continuous illumination of the modified BLM with ultraviolet ($\lambda = 365$ nm) and visible ($\lambda = 450$ nm) light.

photocurrent, $|\Delta I_{max}|$, against λ) was compared with the absorption spectra of 8A5 in CHCl_3 solution (Fig. 3). *Trans*-*cis* isomerization, caused by the ultraviolet light, resulted in a broad band near $\lambda = 440$ nm, which is characteristic of *cis*-8A5. Visible light caused *cis*-*trans* isomerization, as was evident from the reappearance of an intense band at $\lambda = 350$ nm, characteristic of *trans*-8A5. Good agreement between the action spectrum and the absorption spectrum of *cis*-8A5 would suggest that the transient photocurrent in the negative direction is induced by isomerization of *cis*-8A5. It is likely that the photoelectric effect is closely related to the low phase-transition temperature of egg lecithin⁷ and photoisomerization of 8A5. Heating effects caused by light absorption were ruled out as a cause of the observed transients because the electric properties and photoresponse did not depend on temperature for 15–25 °C.

We have carried out preliminary measurements of alternating current across the BLM cell. A sinusoidal wave of amplitude 10 mV and frequency $f = 1$ kHz was applied to the BLM cell. The effective value of the amplitude of the alternating current, ΔA , and the phase difference, $\Delta\phi$, between the a.c. current and input a.c. voltage were measured with a lock-in amplifier. $\Delta\phi$ was decreased by ultraviolet irradiation (time required, ~ 60 s) and returned to the original value by visible light (time required, ~ 30 s) (Fig. 4). On the other hand, ΔA was essentially unchanged by the alternation of illumination. It is reasonable to assume that capacitance of the modified BLM, C_m , is increased by ultraviolet irradiation (8A5 in *cis* form) and returned to its original value by visible light (8A5 in *trans* form), whereas the change in the membrane resistance, R_m , is small. Judging from results on photochemical control of potassium ion permeation across the membrane of the vesicles containing 8A5 (H.F. and Y.Y., manuscript in preparation), the origin of direct current flowing is ion (for example, K^+ , Cl^-) migration through the BLM.

As a simple equivalent circuit of the BLM cell, we have used a combination of the resistance of the aqueous phase, R_s , and the complex impedance of the BLM itself, Z_m , in series. Here Z_m depends on R_m , C_m and f . Our preliminary examination of the frequency dependence of the a.c. impedance ($f = 100$ Hz–10 kHz) with the lock-in amplifier yielded a semicircular plot of the imaginary against the real part of the complex capacitance of the BLM cell under irradiation with visible light⁸. From this plot, we estimated $C_m \approx 10$ nF and $R_s \approx 8 \times 10^4 \Omega$. A model calculation of the transient current in the equivalent circuit showed that the general features of the photoresponse in Fig. 2 are reproduced by the suitable choice of parameters: $R_s = 10^5 \Omega$, $R_m = 10^{10} \Omega$, $C_m = 10$ nF (for 8A5 in *trans* form) or 10.5 nF (for 8A5 in *cis* form), together with time constants of 5–10 s for the decrease in C_m and 30–50 s for the increase in C_m . Time constants employed here are close to the half lives of isomerization by visible light and by ultraviolet light, respectively (H.F. and Y.Y., manuscript in preparation). The change in C_m could be caused by the variation of either the dielectric constant or the thickness of the membrane.

Although further study is needed to establish the detailed relation between molecular and electronic properties, the photochemically induced change of C_m in the modified BLM may be of interest for the development of artificial visual systems. The photoelectric response observed here may bear analogies to that at a molecular level in some biological systems. □

Received 21 January; accepted 30 April 1991.

1. Liu, Z. F., Hashimoto, K. & Fujishima, A. *Nature* **347**, 658–660 (1990).
2. Kano, K., Tanaka, Y., Ogawa, T., Shimomura, M. & Kunitake, T. *Photochem. Photobiol.* **34**, 323–329 (1981).
3. Okahata, Y. *Acc. Chem. Res.* **19**, 57–63 (1986).
4. Tazuke, S., Kurihara, S., Yamaguchi, H. & Ikeda, T. *J. phys. Chem.* **91**, 249–251 (1987).
5. Kumar, G. S., & Neckers, D. C. *Chem. Rev.* **89**, 1915–1925 (1989).
6. Hanai, T., Kajiyama, S., Morita, S. & Koizumi, N. *Bull. Inst. Chem. Res. Kyoto Univ.* **47**, 327–339 (1969).
7. Chapman, D. In *Form and Function of Phospholipids* (eds Ansell, G. B., Hawthorne, J. N. & Dawson, R. M. C.) 117–142 (Elsevier, Amsterdam, 1973).
8. Hanai, T., Haydon, D. A. & Taylor, J. *Proc. R. Soc. A* **281**, 377–391 (1964).
9. Mueller, P., Rudin, D. O., Ti Tien, H. & Wescott, W. C. *Nature* **194**, 979–980 (1962).
10. Mueller, P., Rudin, D. O., Ti Tien, H. & Wescott, W. C. *J. phys. Chem.* **67**, 534–535 (1963).
11. Parker, C. A. *Proc. Roy. Soc. A* **220**, 104–116 (1953).

Decrease of summer tropospheric ozone concentrations in Antarctica

R. C. Schnell*, S. C. Liu†, S. J. Oltmans‡, R. S. Stone*, D. J. Hofmann‡, E. G. Dutton‡, T. Deshler§, W. T. Sturges*, J. W. Harder*, S. D. Sewell||, M. Trainer*† and J. M. Harris‡

* Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado 80309, USA

† NOAA/Aeronomy Laboratory, 325 Broadway, Boulder, Colorado 80303, USA

‡ NOAA/Climate Monitoring and Diagnostics Laboratory, 325 Broadway, Boulder, Colorado 80303, USA

§ Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming 82071, USA

|| Geography Department, University of Colorado, Boulder, Colorado 80303, USA

AS an oxidant and a precursor for other highly reactive oxidants, ozone plays an important role in tropospheric photochemistry. In the upper troposphere, ozone absorbs infrared radiation and is thus an effective greenhouse gas¹. Here we show that surface ozone concentrations at the South Pole in the austral summer decreased by 17% over the period 1976–90. Over the same period, solar irradiance at the South Pole in January and February decreased by 7% as a result of a 25% increase in cloudiness. We suggest that the trend in the summer ozone concentrations is caused by enhanced photochemical destruction of ozone in the lower troposphere caused by the increased penetration of ultraviolet radiation associated with stratospheric ozone depletion, coupled with enhanced transport of ozone-poor marine air from lower latitudes to the South Pole.

According to current theory of the tropospheric ozone budget, photochemical production and stratospheric flux are principal sources of ozone, which are balanced by photochemical destruction and surface deposition^{2–5}. The photochemical production of ozone is mainly limited by the availability of nitrogen oxides, NO_x . In areas where the concentration of NO_x is relatively high, there is net photochemical ozone production^{6,7}. In areas where the concentration of NO_x is low, such as in Antarctica, photochemical destruction dominates^{8,9}, and results in an annual tropospheric ozone cycle with a pronounced summer minimum (Fig. 1).

Figure 1 indicates that as the Sun rises at the pole in late September, ozone begins a period of decline and greater variability. This continues until austral sunset, after which it begins

a steady recovery, peaking in July. This inverse relationship between surface ozone concentration and Antarctic solar radiation suggests that ozone is photochemically destroyed in summer, a process also observed in the Arctic^{10–13}.

At the South Pole, the seasonal decline in ozone begins around the middle of September, but solar radiation levels sufficient to initiate photolytic destruction are not reached until mid-October. Sufficient solar radiation is, however, received at the Antarctic Circle by early September. This suggests that the onset of ozone depletion observed at the South Pole in early spring has its origin at lower latitudes and that ozone-depleted air is subsequently transported towards the pole during transient weather disturbances. For instance, in December 1987 ozone decreased from 39 parts per 10^9 (p.p.b.v.) to 17 p.p.b.v. in an hour when a warm marine air mass arrived. Such transport events are common occurrences at South Pole^{14–17}. Isobaric back trajectories¹⁸ on the 500-mbar surface indicate that air transport often occurs from 65°S to the South Pole within three days. For comparison, the photochemical lifetime of ozone in the lower troposphere in the austral spring is about 40 days at 65°S.

An analysis of the South Pole ozone record from 1975 to 1989

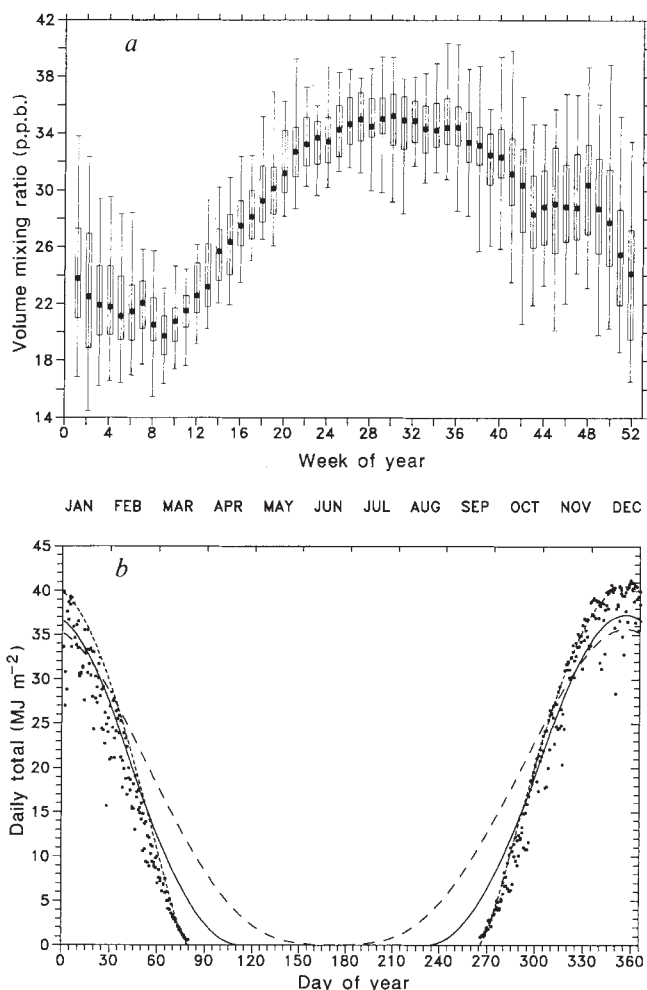


FIG. 1 *a*, Seasonal mean ozone cycle and the variability at the South Pole (2.8 km above mean sea level) averaged weekly for January 1975 to December 1989, as measured by the National Oceanic and Atmospheric Administration's Climate Monitoring and Diagnostics Laboratory. The median (dot), the upper and lower quartiles (ends of the boxes), and the upper and lower fifth percentiles (ends of the vertical lines) of weekly average ozone values are shown. *b*, Observed daily total global short-wave radiation at the South Pole for 1986 and 1987 (dots) and modelled clear-sky solar radiation at McMurdo, 77.9° S (—), at the latitude of the Antarctic Circle (— — —) and at the South Pole (— — — — —).

(Fig. 2) shows a statistically significant downward trend of $\sim 17\%$ in the near-surface ozone concentration for the austral summer months (averaged for December, January and February) over the 14-year period (rate of ~ 0.32 p.p.b.v. per year when fitted linearly). No statistically significant trend can be found for other seasons, although a tendency towards a small decrease in ozone begins in October. The results suggest that since 1976, there has been an amplification of the normal summertime ozone destruction illustrated in Fig. 1. Presumably, photochemical processes are important both in the seasonal cycle and in the longer term decrease in surface ozone at the South Pole.

Possible contributors to this downward trend of surface ozone for austral summer at the South Pole are (1), a decrease of the stratospheric ozone flux; (2), an ozone-destroying reaction with bromine, as observed in the Arctic^{10,11}; (3), an increase in transport of ozone-depleted air from lower latitudes; and (4), enhanced photochemical ozone destruction due to greater solar ultraviolet flux as a result of the stratospheric ozone hole.

Intuitively, the first mechanism is attractive because there is clear evidence for large ozone loss in the lower stratosphere in austral spring^{19,20}. Ozone profiles, obtained during the formation of the Antarctic ozone hole¹⁹⁻²² (August to early November) since 1986 at McMurdo Station, indicate that the spring decrease in tropospheric ozone extends upward until a stable layer restricts vertical mixing; this usually is the tropopause at McMurdo. This suggests that the reduction in tropospheric ozone originates in the lower troposphere. In addition, Fig. 3 shows that tropospheric ozone begins to decline after the total column ozone has reached a minimum in mid-October (about day 290); the decline begins earlier and proceeds

faster at low altitudes. If a reduced stratospheric ozone flux into the troposphere, associated with the ozone hole, were primarily responsible for the tropospheric ozone loss, one would expect to see the decline begin earlier at all altitudes and affect the lower altitudes last.

The second mechanism, ozone destruction by reaction with bromine¹⁰, is also unlikely to be a main cause of the downward trend because concentrations of biogenic bromine gases observed at McMurdo are small¹³, and the seasonal ozone cycle

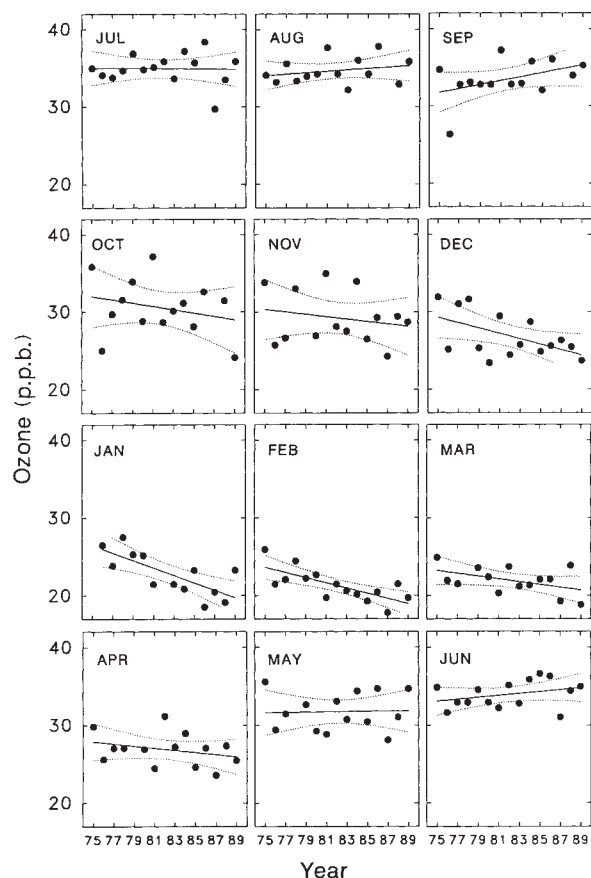


FIG. 2 Mean mixing ratio for surface ozone at the South Pole for each month during 1975-89. The solid line is a linear least-squares fit to the data for each month. The dotted lines are 95% confidence intervals for the trend lines. The December, January, February and March trends are significant at the 95% level.

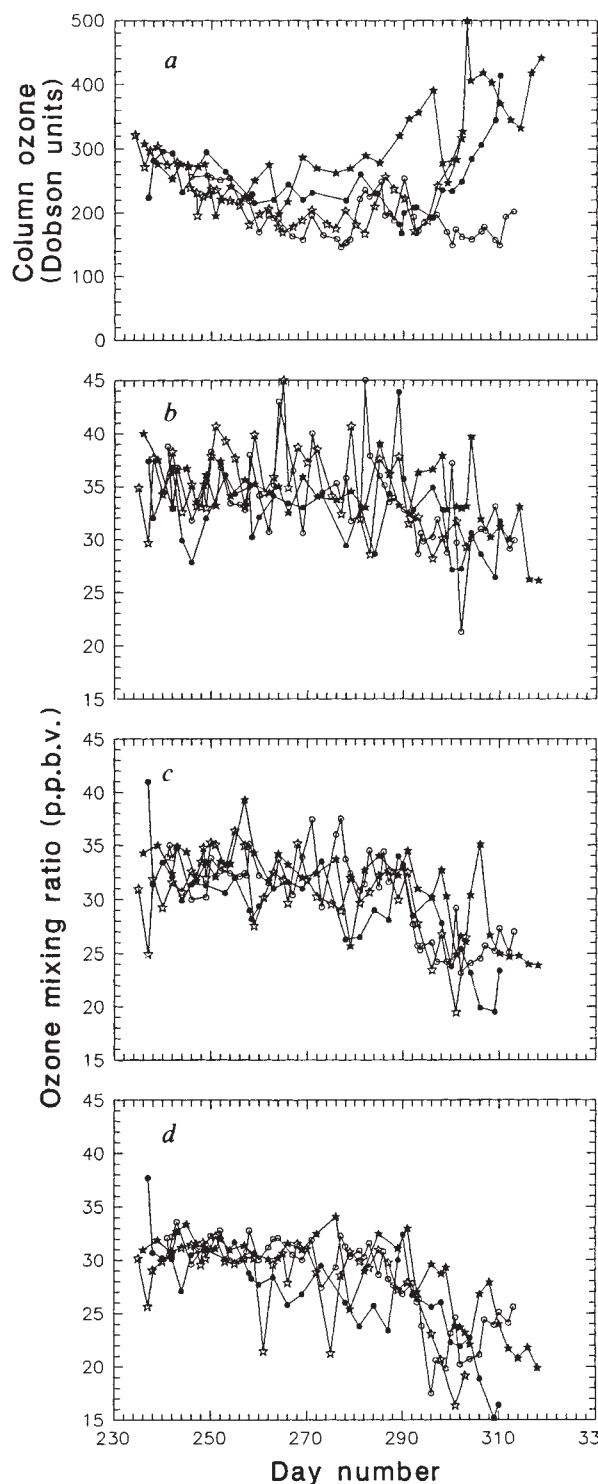
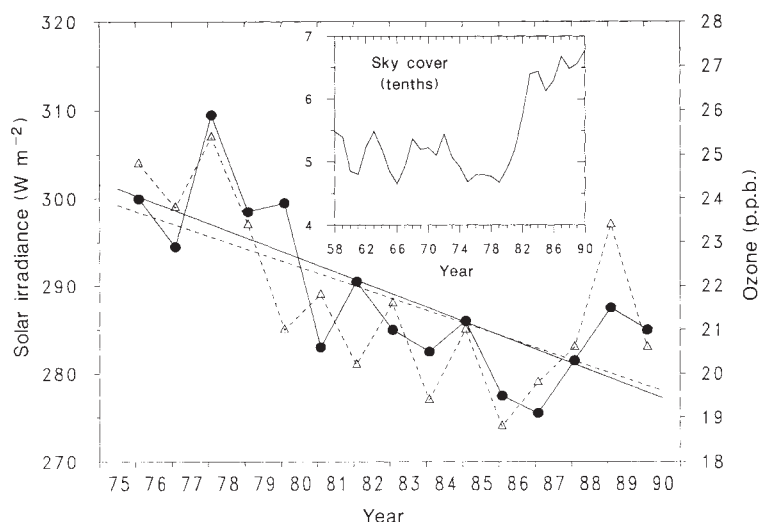


FIG. 3 a, Total column ozone from integration of ozonesonde profiles. b-d, Ozone mixing ratio for 0.5-km averages centred at b, 5 km, c, 3 km and d, 1 km at McMurdo Station, Antarctica, for 1986-89.

FIG. 4 a, Mean surface ozone concentrations and linear trend in January and February (dots and solid line, referenced to right scale) compared with the mean January–February solar irradiance flux²³ and linear trend (triangles and dashed lines, referenced to left scale), measured at the South Pole from 1976 to 1990. The data points are plotted for 31 January of each year. Inset, long-term (1958–90) record of January–February sky cover at the South Pole.



at the South Pole does not have the bimodal structure observed in the Arctic¹¹.

Evidence of enhanced transport to the South Pole (the third mechanism) is supported by solar irradiance measurements made at the surface at the South Pole in January and February that show statistically significant downward trends anticorrelated with the observed concurrent sky cover²³. The downward trend in the mean January–February solar irradiance is well correlated with a downward trend in surface ozone concentrations measured at the pole (Fig. 4). Shown in the inset of Fig. 4 is the observed January–February sky cover for the South Pole from 1958 to 1990. Cloudiness increased by 25% over the last decade of the record. These data suggest that enhanced transport of ozone-depleted air (due to photochemical ozone destruction) from lower latitudes contributes to the surface ozone loss, at least during the late austral summer months at the South Pole. But an examination of the solar irradiance data by month shows that this significant downward trend in solar irradiance exists only in January and February, whereas the decrease of surface ozone seems to start as early as October and last until the end of April (Fig. 2). Thus, although the transport mechanism probably contributes to the observed downward trend in tropospheric ozone during January and February, it may play a minor part in the other months.

The fourth mechanism, enhanced photochemical destruction caused by increased ultraviolet flux resulting from the depletion of ozone in the stratosphere, is likely to be a principal contributor to tropospheric ozone depletion in Antarctica, at least during spring. Table 1 lists the rate of ozone loss due to the photochemical mechanism^{8,9}, calculated from a model^{6,24}, for 11–31 October. Because the ozone depletion starts at lower latitudes,

we have made model calculations for 75°S and 65°S. The ambient conditions assumed for 75°S were 30 p.p.b.v. O₃, 60 p.p.b.v. CO, 1,600 p.p.b.v. CH₄, 1 p.p.t.v. (part per 10¹² volume) NO_x, –15 °C surface temperature with a lapse rate of 6 °C per km and a relative humidity of 80%. These values are consistent with available observations near Antarctic areas^{25–27}. The same conditions were assumed for 65°S, except that the surface temperature was set to –8 °C. Clear-sky conditions were assumed, and an albedo of 0.85 was adopted²⁸. Overhead ozone column density for this model was fixed at 200 and 250 Dobson units for 75°S and 65°S, respectively (consistent with average values between 1986 and 1989^{29,30}), to compare calculated results with the observations depicted in Fig. 3.

The observed ozone depletion rate at McMurdo (77.9°S) in the years 1986–89 was ~0.4, 0.3 and 0.2 p.p.b.v. per day near day 300 at altitudes of 1, 3 and 5 km, respectively. For comparison, the calculated ozone loss rate at 75°S is 0.43, 0.28 and 0.12 p.p.b.v. at the corresponding altitudes (Table 1). The rates calculated for 65°S are ~65% greater than those at 75°S. Thus, the observed ozone depletion above McMurdo can easily be accounted for by the classical photochemical loss mechanism. For comparison, the largest surface ozone depletion rate at the South Pole occurs between days 250 and 300 and is ~0.12 p.p.b.v. per day.

Further model calculations show that the decrease of 17% in surface ozone observed at the South Pole between 1976 and 1989 during austral summer would require a corresponding decrease of ~10% in the overhead ozone column in the spring and summer. This compares favourably with the observed values. For instance, at the South Pole the 10-year-average column ozone densities for October, November and December 1980–89 are 29%, 21% and 10% less, respectively, than the averages for 1970–79³¹. At several locations on the coast of Antarctica, overhead column ozone densities have decreased by similar percentages³² and data from satellite observations show that these column ozone reductions extend to north of the Antarctic Circle^{29,33}.

Amplification of the photochemical destruction of tropospheric ozone in Antarctica during austral springtime can thus be explained through the increased penetration of ultraviolet radiation associated with the stratospheric ozone hole. There is no obvious explanation, however, for the marked changes in the summertime poleward transport of marine air beginning in about 1980. Intuitively, the driving force is expected to be in the troposphere. Anomalies in sea surface temperatures in the southern hemisphere have been found to be correlated with the overhead ozone variations at the South Pole^{34–36}. Such anomalies may also affect the strength of the equator-to-pole circulation³⁴

TABLE 1 Model-calculated ozone loss rates compared with observed ozone depletion rates

	Overhead O ₃ (Dobson units)	Ozone loss rate (p.p.b.v. per day)		
		1 km	3 km	5 km
Calculated				
75°S	200	0.43	0.28	0.12
65°S	250	0.71	0.5	0.22
75°S	267	0.31	0.2	0.08
Observed McMurdo	200 ± 50	0.4	0.3	0.2

Calculated rates are for two latitudes and three fixed overhead ozone column densities in the Antarctic atmosphere during 11–31 October. Observed depletion rates are for McMurdo (77.9°S) near day 300 (25 October) at 1, 3 and 5 km.

and thus affect cloudiness in the polar region. We are pursuing research in this area. □

Received 31 January; accepted 9 May 1991.

1. Ramanathan, V., Cicerone, R. S., Singh, H. B. & Kiehl, J. T. *J. geophys. Res.* **90**, 5547–5566 (1985).
2. Crutzen, P. J. in *Physics and Chemistry of the Upper Atmosphere* (ed. McCormac, B.) 389 (Reidel, Dordrecht, 1973).
3. Chameides, W. L. & Walter, J. C. G. *J. geophys. Res.* **78**, 8781–8760 (1973).
4. Fishman, J., Solomon, S. & Crutzen, P. J. *Tellus* **31**, 432–446 (1979).
5. Liu, S. C., Kley, K., McFarland, M., Mahlman, J. D. & Levy, H. II. *J. geophys. Res.* **85**, 7546–7552 (1980).
6. Liu, S. C. *et al.* *J. geophys. Res.* **92**, 4191–4207 (1987).
7. Fishman, J., Watson, C. E., Larsen, J. C. & Logan, J. A. *J. geophys. Res.* **85**, 3599–3618 (1990).
8. Liu, S. C., McFarland, M., Kley, D., Zafiriou, O. & Huebert, B. *J. geophys. Res.* **88**, 1360–1368 (1983).
9. Chameides, W. L. *et al.* *J. geophys. Res.* **92**, 2131–2152 (1987).
10. Barrie, L. A., Bottenheim, J. W., Schnell, R. C., Crutzen, J. & Rasmussen, R. *Nature* **334**, 138–141 (1988).
11. Oltmans, S. J. *et al.* *Atmos. Envir.* **23**, 2431–2441 (1989).
12. Finlayson-Pitts, B. J., Livingston, F. E. & Berko, H. N. *Nature* **343**, 622–625 (1990).
13. Sturges, W. T., Schnell, R. C., Landsberger, S. & Harris, J. M. *Atmos. Envir.* (submitted).
14. Ohtake, T. *Antarct. J. U.S.* **13**, 174–175 (1978).
15. Bodhaine, B. A., DeLuise, J. J. & Harris, J. M. *Tellus* **B38**, 223–235 (1986).
16. Stone, R. S., Dutton, E. G. & DeLuise, J. J. *Antarct. J. U.S.* **24–5**, 230–232 (1990).
17. Stone, R. S. & Kahl, J. D. *J. geophys. Res.* (in the press).
18. Harris, J. M. *NOAA Tech. Memo. ERL ARL-116* (NOAA Environmental Research Laboratories, Boulder, Colorado, 1982).
19. Hofmann, D. J., Harder, J. W., Rolf, S. R. & Rosen, J. M. *Nature* **326**, 59–62 (1987).
20. Hofmann, D. J., Harder, J. W., Rosen, J. M., Hereford, J. V. & Carpenter, J. R. *J. geophys. Res.* **94**, 16527–16536 (1989).
21. Desher, T., Hofmann, D. J., Hereford, J. V. & Sutter, C. B. *Geophys. Res. Lett.* **17**, 151–154 (1990).
22. Desher, T., Hofmann, D. J. & Hereford, J. V. *J. geophys. Res.* **95**, 10023–10035 (1990).
23. Dutton, E. G., Stone, R. S., Nelson, D. W. & Mendonca, B. G. *J. Clim.* (in the press).
24. Liu, S. C. & Trainer, M. *J. Atmos. Chem.* **6**, 221–234 (1988).
25. Schwerdtfeger, W. *Weather and Climate of the Antarctic* (Elsevier, Amsterdam, 1984).
26. Seiler, W. & Fishman, D. J. *J. geophys. Res.* **86**, 7255–7266 (1981).
27. Drummond, J. W., Ehrl, D. H. & Volz, A. *J. geophys. Res.* **93**, 15831–15849 (1988).
28. Warren, S. G. *Rev. Geophys. Space Sci.* **20**, 677–89 (1982).
29. WMO, *Scientific Assessment of Stratospheric Ozone: Global Ozone Research and Monitoring Project Rep. No. 20* (World Meteorological Organization, Geneva, 1989).
30. Stolarski, R. S., Krueger, A. J. & Schoeberl, M. R. in *Proc. Quad. Ozone Symp. Gottingen* 185–188 (Deepak, Hampton, Virginia, 1990).
31. Komhyr, W. D., Oltmans, S. J., Grass, R. D., Evans, R. D. & Leonard, R. K. in *Geophysical Monitoring for Climatic Change, No. 17: Summary Rep. 1988* (eds Elkins, J. W. & Ross, R. M.) 36–42 (National Oceanic and Atmospheric Administration, Boulder, Colorado, 1989).
32. Bojkov, R. D. *Geophys. Res. Lett.* **13**, 1236–1239 (1986); updated in WMO Fact Sheet *The WMO Global Ozone Observing System* (World Meteorological Organization, Geneva, 1990).
33. Stolarski, R. S., Krueger, A. J. & Schoeberl, M. R. in *Proc. Quad. Ozone Symp. Gottingen* 185–188 (A. Deepak, Hampton, Virginia, 1990).
34. *UK Stratospheric Ozone Review Group Rep. No. 1* (HMSO, London, 1987).
35. Komhyr, W. D., Oltmans, S. J., Grass, R. D. & Leonard, R. K. *Can. J. Phys.* (in the press).
36. Angell, J. K. *Geophys. Res. Lett.* **17**, 1569–1572 (1990).

ACKNOWLEDGEMENTS. We thank A. Tuck for helpful discussions. This research was supported in part by the NASA Upper Atmosphere Research Program, the Office of Naval Research, University Research Initiative, the NSF Division of Polar Programs and the NOAA Radiatively Important Trace Species component of the Climate and Global Change Program.

Rapid transitions of the ocean's deep circulation induced by changes in surface water fluxes

Thomas F. Stocker* & Daniel G. Wright†

Centre for Climate and Global Change Research, McGill University, Montreal, Quebec, Canada H3A 2K6

* Now at: Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA

† Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada B2Y 4A2

DEEP water in the world's oceans flows predominantly from the northern North Atlantic into the Pacific¹, slowly upwells on the way to become part of the upper warm-water circulation, and returns to the North Atlantic. The stability of this thermohaline conveyor belt has recently been questioned on the basis of palaeoclimatic data from deep-sea sediment and ice cores^{2,3}. Different modes of deep circulation have been confirmed in numerical ocean models^{4–6}, and the present-day circulation has been shown to be sensitive to changes in the surface-water budget⁷. Here we use an idealized model^{4,5} to examine the hypothesis that small changes in the atmospheric flux of fresh water from the Atlantic to the Pacific could force the thermohaline circulation to switch

between two stable modes. Our results indicate that a decrease of this flux can reverse the Atlantic circulation, although the Pacific thermohaline circulation does not change direction. This is consistent with reconstructions of conditions in the Atlantic Ocean during the last glacial obtained from deep-sea cores⁸. To re-establish the conveyor belt, the fresh-water flux need be increased only slightly beyond its present value.

Quaternary ice ages have been related to interactions between ice cover, lithosphere and atmosphere responding to astronomically forced variations in seasonality of insolation⁹. Isotope measurements in deep-sea cores indicate that at the last glacial maximum, 18,000 BP (before present), the deep ocean had properties different from today^{8,10}; deep-ocean temperatures were cooler¹¹, North Atlantic deep-water production was reduced⁸ and considerably more bottom water was flowing from the Southern Ocean into the Atlantic¹⁰. Several authors^{2,3} conclude that the ocean may have been in a different circulation mode than the present one. Broecker and Denton² suggested that changes in the hydrological cycle could induce mode changes in the ocean circulation with consequent rapid climate changes.

Recent results of two- and three-dimensional numerical ocean models^{5,6} are consistent with the original idea¹² that different modes of operation of the global thermohaline circulation may exist. It was also shown that the present-day meridional circulation is sensitive to anomalous fresh-water fluxes applied locally to the surface of the Atlantic⁷, and to changes in the net Atlantic-to-Pacific fresh-water flux, F , through the atmosphere⁵. The present estimate for F lies in the range of +0.1 to +0.45 Sv (1 Sv = 10^6 m³ s⁻¹)^{2,13}. Model simulations of atmospheric conditions at the last glacial maximum corroborate the possibility of glacial-to-interglacial changes in the hydrological cycle and suggest that F was greatly reduced or even negative¹⁴.

To investigate the implications of such changes in the thermohaline circulation, we used an idealized ocean model consisting of two coupled, zonally averaged basins representing the Atlantic and Pacific^{4,5}. The model is purely buoyancy-driven by surface fluxes of heat and salt, and it yields results consistent with more complex three-dimensional general circulation models for the oceans^{6,15}. After initial integration to steady state by relaxing surface temperature and salinity to prescribed profiles, we apply mixed boundary conditions, whereby the sea surface temperature is still relaxed and the surface salt flux is kept fixed.

The steady state obtained under present-day surface forcing (mode A) shows a conveyor-belt circulation^{1,5}: deep water is formed in the North Atlantic and flows into the Pacific where broad upwelling occurs (Fig. 1a). Temperature and salinity fields in both Pacific and Atlantic compare favourably with observed zonal averages (Fig. 1b and c)¹⁶. Associated with this conveyor belt is a net Atlantic-to-Pacific fresh-water flux through the atmosphere, F , of 0.3 Sv, which is close to the estimate of 0.45 Sv obtained by Baumgartner and Reichel¹³. The circulation is not sensitive to changes in F unless it is reduced to ~0.03 Sv, at which point the system changes to a second equilibrium (mode B). In this state the conveyor belt is shut off, and the Atlantic circulation resembles that of the Pacific (Fig. 2a); deep water is now only formed in the Southern Ocean. As a consequence, the temperature and particularly the salinity distribution in the two ocean basins have changed (Fig. 2b and c). Just as mode A compares well with today's circulation pattern, mode B is consistent with the reconstruction by Duplessy *et al.*¹⁰ of ocean conditions during the last glacial maximum.

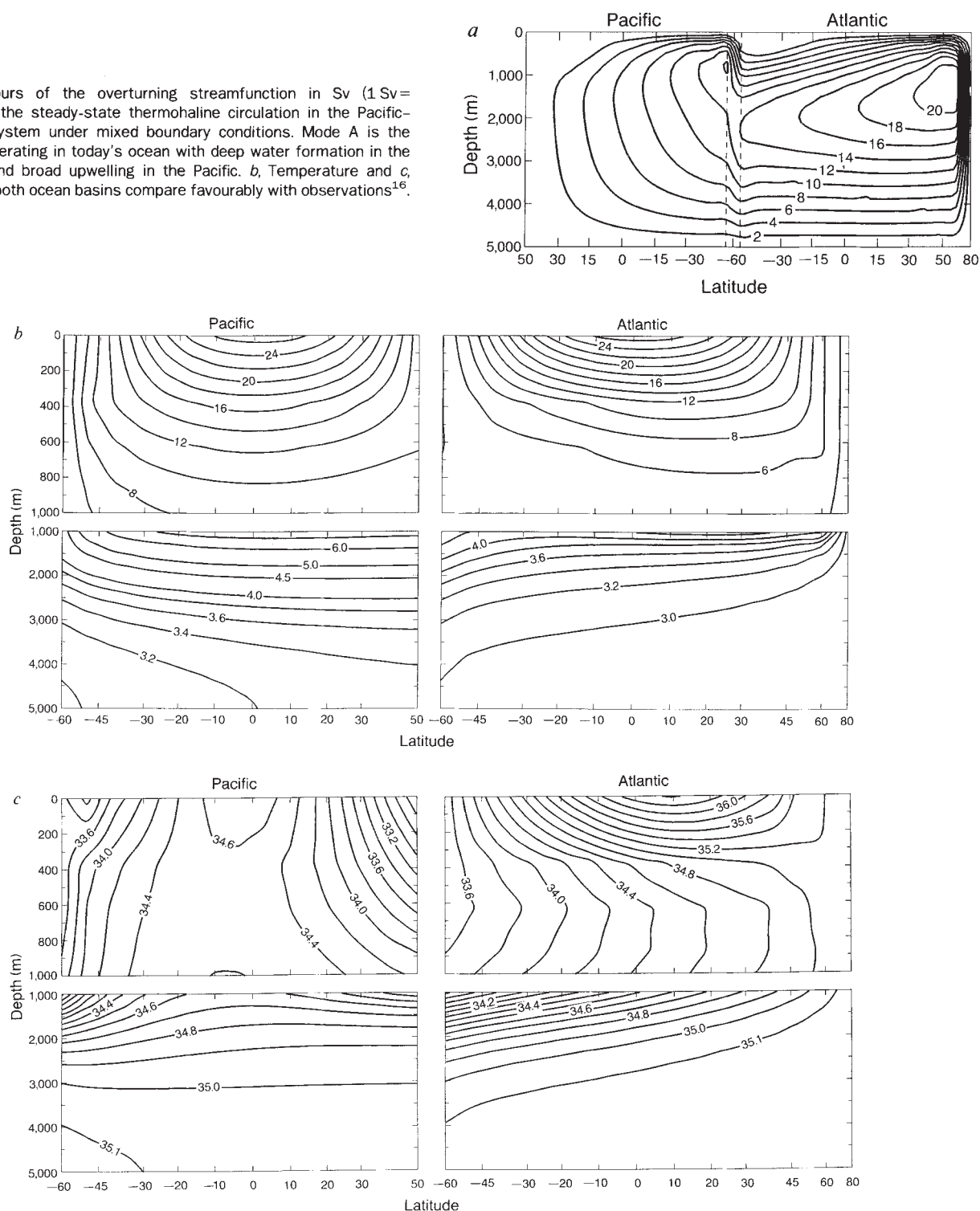
Once in mode B, the Atlantic-to-Pacific fresh-water flux can be increased to its initial value without significant change of the thermohaline circulation, confirming the existence of two stable modes under identical boundary conditions¹². But if a threshold of ~0.36 Sv (1.2 times the original value) is exceeded, the conveyor belt is re-established. This behaviour is qualitatively consistent with the observation that the climate system remains

stable for many millennia, then relatively quickly undergoes a transition to a new state where it again remains for many millennia². Figure 3 shows the meridional heat transport carried northward by the Atlantic as a function of F_1 , exhibiting the hysteresis behaviour described above. Mode A on the upper branch of the hysteresis curve is associated with a maximum northward heat flux of $\sim 0.6 \times 10^{15}$ W, whereas a similar heat flux is directed southward in mode B (lower branch). Shutting off the ocean heat flux carried by the Atlantic could cause a climatic event similar to the Younger Dryas³: reversing this heat flux would certainly have a considerable global impact.

Decreasing the atmospheric Atlantic-to-Pacific fresh-water

flux thus alters the global circulation from the conveyor belt to a flow with little exchange between the two ocean basins. Equally important, an increase in the Atlantic-to-Pacific fresh-water flux is a plausible and consistent mechanism by which the present-day temperature and salinity distributions in the Pacific and Atlantic ocean could be re-established and the conveyor-belt circulation resumed. Several experiments with varying resolution, vertical diffusivities and relaxation temperatures show that the locations of the transition points on the hysteresis curve are sensitive to these modifications, but the circulation, temperature and salinity patterns of the stable modes, as well as the existence of hysteresis, are robust features.

FIG. 1 *a*, Contours of the overturning streamfunction in Sv ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$) for the steady-state thermohaline circulation in the Pacific-Atlantic basin system under mixed boundary conditions. Mode A is the conveyor belt operating in today's ocean with deep water formation in the North Atlantic and broad upwelling in the Pacific. *b*, Temperature and *c*, salinity fields in both ocean basins compare favourably with observations¹⁶.



The cessation of the conveyor belt and its subsequent resumption are illustrated in Fig. 4, where time series of the mean sea-surface salinity for Atlantic and Pacific are given. Both transitions are triggered by modifying F . The experiment starts from mode A with F abruptly reduced to 0 Sv. During the first few hundred years, reduced Atlantic evaporation causes a steady increase of deep-water formation in the South Atlantic and a reduction of the Atlantic surface salinity to the Pacific level. The subsequent transition to mode B is apparent in the rapid drop of Atlantic surface salinity, leading to a reversed surface salinity contrast between Atlantic and Pacific. At $t = 5,000$ yr, F is increased from 0 Sv to 0.6 Sv, triggering the resumption of

the conveyor belt. Enhanced fresh-water loss leads to maximum surface salinity in the Atlantic. Comparison of Figs 1c and 2c shows that mode changes involve considerable rearrangement of the salinity distribution in both ocean basins.

Several features of Quaternary climate evolution, such as the rapid termination and resumption of the present-day conveyor belt under steady forcing, cessation of North-Atlantic deep-water formation and advection of Southern Ocean deep water into the Atlantic during the glacial, are reproduced by this idealized model. This suggests the importance of the inter-ocean transport of fresh water and indicates that the present-day contrast between surface salinity of the Atlantic and Pacific was

FIG. 2 *a*, Contours of the overturning streamfunction for the second steady state under the same mixed boundary conditions. In mode B the Atlantic flow is reversed and deep water is formed only in the Southern Ocean, whereas the thermohaline circulation in the Pacific remains unchanged. This mode could have operated during much of the glacial. Although the temperature field, *b*, is only slightly altered, the salinity distribution in both basins, *c*, is substantially different from mode A.

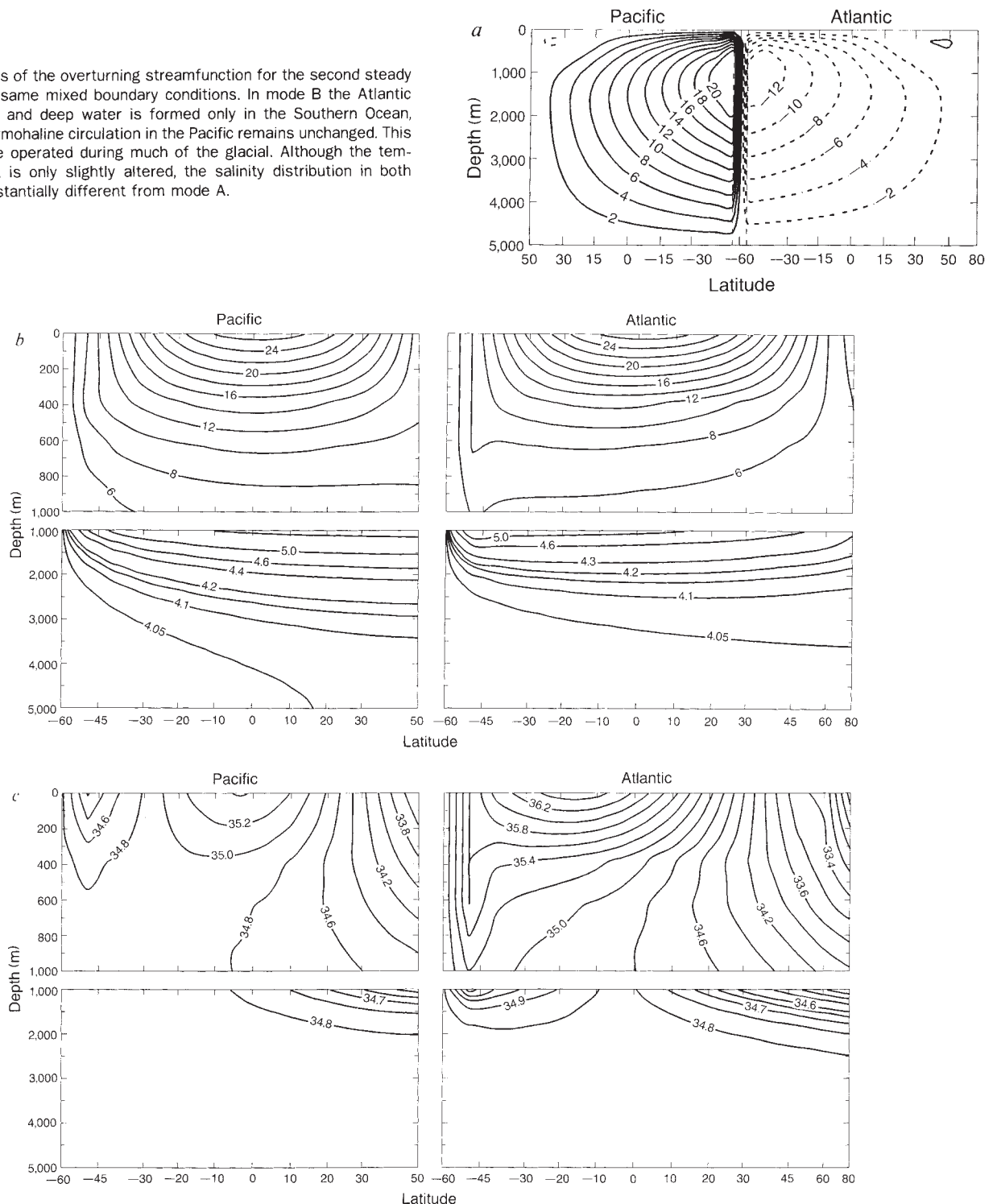


FIG. 3 Hysteresis characteristic of the two stable equilibria for changes of the atmospheric Atlantic-to-Pacific fresh-water transport. The star on the upper branch (mode A) indicates the present-day state according to our model; its location on the curve depends on the model parameters. Results from a three-dimensional model⁷ indicate that the present-day climate may be much closer to the transition point and hence more sensitive to anomalies in the fresh-water flux. Mode A becomes unstable if the Atlantic evaporation excess is reduced to less than 0.03 Sv, and a rapid transition to mode B on the lower branch occurs. Mode B is unstable for an Atlantic evaporation excess larger than 0.36 Sv, and the system returns to mode A. Variation of the atmospheric Atlantic-to-Pacific fresh-water transport is a plausible mechanism by which the global conveyor belt could be shut down as well as re-established.

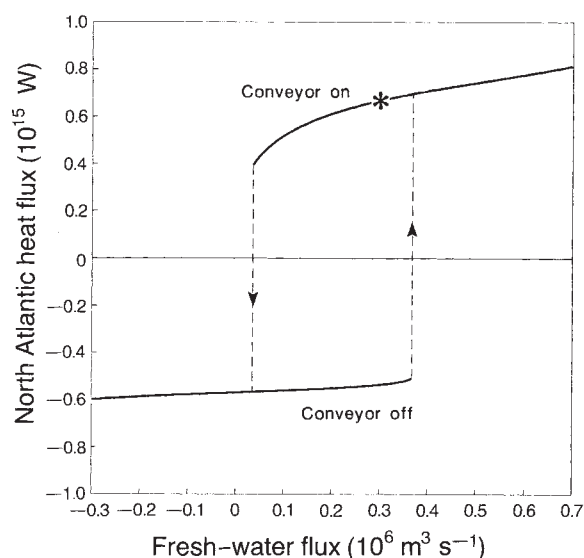
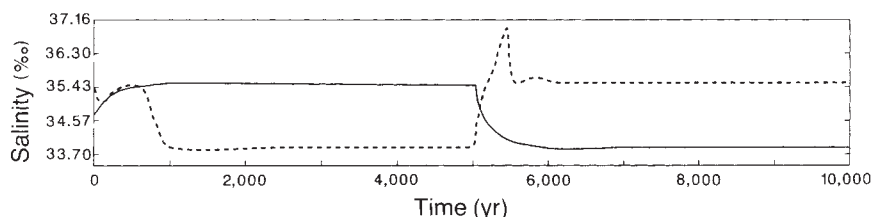


FIG. 4 Evolution during mode changes of the mean surface salinity (parts per 10³) in the Pacific (solid) and Atlantic (dashed). The transition to mode B is initiated at $t=0$ yr by reducing the net Atlantic-to-Pacific fresh-water transport to zero. The transition back to the conveyor belt, mode A, is triggered at $t=5,000$ yr by an increase in fresh-water transport to 0.6 Sv. The transfer of fresh water re-establishes maximum surface salinity in the Atlantic. The transition to mode B is a relatively smooth process with a steady decrease of North Atlantic deep water production, but the resumption of the conveyor belt involves strong convective overturning in the North Atlantic in the early stages.



probably absent or even reversed in the glacial ocean. Small and realistic changes of the atmospheric water balance provide a plausible mechanism by which fundamental mode changes of the global thermohaline circulation (both shut-down and resumption of the conveyor belt) can be induced. This supports the earlier hypothesis, based on the reconstruction of various palaeoclimatic records^{2,3}, that the world ocean has more than one stable mode of operation. But rather than the inverse conveyor belt³, whereby the circulation reverses in both ocean basins, the model indicates that changes of the thermohaline flow should involve primarily the Atlantic Ocean, while the deep circulation in the Pacific changes little.

Changes to the surface-water budget in the Atlantic and Pacific oceans may thus cause reversals of the Atlantic thermohaline flow. Because of the hysteresis, possible mode transitions depend critically on the initial state and the direction of the change in the hydrological cycle. Determining the present state and predicting future changes in the hydrological cycle are therefore central issues for climate research. Recent results suggest that F may increase in a climate with twice the present carbon dioxide concentration¹⁷. From the findings presented here, it seems unlikely that this will result in fundamental changes in ocean mode of the type described above. □

Received 21 December 1990; accepted 16 May 1991.

- Gordon, A. L. *J. geophys. Res.* **91**, 5037–5046 (1986).
- Broecker, W. S. & Denton, G. H. *Geochim. cosmochim. Acta* **53**, 2465–2501 (1989).
- Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21–26 (1985).
- Wright, D. G. & Stocker, T. F. *J. phys. Oceanogr.* (in the press).
- Stocker, T. F. & Wright, D. G. *J. phys. Oceanogr.* (in the press).
- Marotzke, J. & Willebrand, J. *J. phys. Oceanogr.* (in the press).
- Maier-Reimer, E. & Mikolajewicz, U. in *Oceanography* (eds Ayala-Castañares, A. et al.) 87–100 (UNAM Press, Mexico, 1989).
- Boyle, E. A. & Keigwin, L. *Nature* **330**, 35–40 (1987).
- Berger, A. *Rev. Geophys.* **26**, 624–657 (1988).
- Duplessy, J. C. et al. *Paleoceanography* **3**, 343–360 (1988).
- Labeyrie, L. D., Duplessy, J. C. & Blanc, P. L. *Nature* **327**, 477–482 (1987).
- Stommel, H. *Tellus* **13**, 224–230 (1961).
- Baumgartner, A. & Reichel, E. *The World Water Balance* (Elsevier, Amsterdam, 1975).
- Miller, J. R., Russell, G. L. *Paleoceanography* **5**, 397–407 (1990).
- Marotzke, J., in *Ocean Circulation Models: Combining Data and Dynamics* (eds Anderson, D. L. T. & Willebrand, J.) 501–511 (Kluwer, Dordrecht, 1988).
- Levitus, S. *Climatological Atlas of the World Ocean* (NOAA Prof. Pap. 13, 1982).
- Zaucker, F. & Broecker, W. S. *J. geophys. Res.* (in the press).

ACKNOWLEDGEMENTS. We are grateful for comments by J. Lazier, J. Loder and B. Petrie. This work was supported by the Swiss NSF and NSERC Canada.

Incorporation of hydroxyl in upper-mantle clinopyroxenes

J. R. Smyth*, D. R. Bell† & G. R. Rossman†

* Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309-0250, USA

† Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

WATER (and hydroxyl, OH) plays an important part in determining the properties of minerals and melts in the Earth's upper mantle. The main hydroxyl-bearing phases found in rocks from the upper mantle, phlogopite and amphibole, are not believed to exist in significant quantities at depth. Some of the less abundant phases found in these rocks contain small amounts of hydrous components^{1–3}, but do not constitute an important reservoir for water. Traces of hydroxyl have been found in common mantle phases^{4–6}, but not at concentrations high enough to account for the amount of water thought to be present at depth. An exception

TABLE 1 Electron microprobe analyses of eclogite clinopyroxenes in wt %

Sample	SBB-1	SBB-5H	SBB-7P	SBB-14	SBB-20	SBB-46	SBB-104	SDC-1	SDC-2	SRV-4	XM-37	HRV-147
SiO ₂	57.14	55.11	55.06	55.05	55.09	55.81	55.31	55.60	55.71	54.59	56.10	56.25
TiO ₂	0.37	0.08	0.06	0.14	0.11	0.28	0.22	0.17	0.17	0.09	0.38	0.30
Al ₂ O ₃	19.08	7.24	2.61	0.31	5.93	10.40	6.27	17.73	17.58	1.25	14.10	14.28
Cr ₂ O ₃	0.03	0.08	3.35	1.82	0.09	0.80	0.03	0.05	0.04	0.08	0.15	0.07
FeO	2.46	3.31	1.74	2.27	4.32	2.44	4.07	1.28	1.27	3.11	2.54	1.98
MnO	0.03	0.03	0.08	0.09	0.04	0.02	0.04	0.01	0.01	0.05	0.02	0.03
MgO	4.14	12.39	15.72	17.46	12.59	10.19	12.72	5.99	6.00	17.11	7.72	8.15
CaO	7.70	17.56	18.93	21.72	17.90	14.53	17.06	10.32	10.37	23.23	11.86	12.75
Na ₂ O	9.12	4.40	2.85	1.31	4.07	5.78	4.29	8.90	8.90	0.66	6.99	6.76
K ₂ O	0.05	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.07	0.29	0.25
P ₂ O ₅	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	n.d
Total	100.12	100.20	100.40	100.17	100.14	100.35	100.01	100.05	100.05	100.24	100.15	100.82
Atoms per 6 oxygens												
Si	1.971	1.973	1.981	1.995	1.986	1.971	1.988	1.933	1.937	1.981	1.966	1.957
Ti	0.010	0.002	0.002	0.004	0.003	0.007	0.006	0.004	0.004	0.003	0.010	0.002
Al	0.776	0.306	0.111	0.013	0.252	0.433	0.266	0.727	0.720	0.054	0.582	0.585
Cr	0.001	0.002	0.095	0.052	0.003	0.022	0.001	0.001	0.001	0.002	0.004	0.008
Fe	0.071	0.099	0.052	0.069	0.130	0.072	0.122	0.037	0.037	0.094	0.074	0.058
Mn	0.001	0.001	0.002	0.003	0.001	0.001	0.001	0.000	0.000	0.002	0.001	0.001
Mg	0.213	0.661	0.843	0.943	0.677	0.537	0.682	0.311	0.311	0.926	0.403	0.423
Ca	0.285	0.674	0.730	0.843	0.691	0.550	0.657	0.385	0.386	0.903	0.445	0.475
Na	0.610	0.305	0.199	0.092	0.284	0.396	0.299	0.600	0.600	0.046	0.475	0.456
K	0.002	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.003	0.013	0.011
P	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Total	3.938	4.024	4.014	4.015	4.027	3.994	4.022	3.998	3.998	4.014	3.974	3.975
Vacancy	0.062	-0.024	-0.014	-0.015	-0.027	0.006	-0.022	0.002	0.002	-0.014	0.026	0.025

is suggested by a report⁷ that the pyroxene (omphacite) in an eclogite nodule from the Roberts Victor kimberlite pipe contains up to 1,000 p.p.m. OH. Here we show that some of this hydroxyl is associated with cation vacancies in sodic clinopyroxene, and that these pyroxenes may be an important reservoir for hydrous components in the upper mantle.

To constrain the hydroxyl contents of minerals in the eclogite suite of rocks, we have measured infrared absorption spectra of a representative suite of clinopyroxenes covering the range of compositions found in mantle eclogites. The aims of this study were to see if hydroxyl is associated with the M2 cation vacancy previously noted in some of the pyroxenes⁸, and to see if the hydroxyl content can help to constrain the origin and evolution of these rocks.

Pyroxene samples were selected from the well documented suite of mantle eclogites from South African kimberlites⁹⁻¹¹. Samples were selected to span the range of pyroxene compositions observed in mantle eclogites, and to have inclusion-free and fracture-free areas of at least 150 μm in their shortest dimension. Eleven samples were selected for the study. Samples denoted SBB are from the Bobbejaan mine; samples denoted SDC are from the Dan-Carl mine (both at Bellsbank); and samples SRV-4, XM-37 and HRV-147 are from the Roberts Victor Mine near Boshof, Orange Free State, South Africa.

Samples SDC-1 and 2 are from corundum-bearing gneiss, sample SBB-1 is from a kyanite eclogite, whereas the rest are from bimineralline eclogites, except for SBB-14 which is a chromium diopside megacryst. Sample HRV-147 is from an inhomogeneous, kyanite-bearing eclogite described by Hatton¹² and used by Skogby *et al.*⁸. All samples were analysed with the electron microprobe. Pyroxene analyses are presented in Table 1 together with atoms per formula unit (12 oxygens). Samples SDC-1 and 2 and SBB-1 are rich in jadeite (Jd), whereas sample SRV-4 is low in sodium and contains no jadeite in the endmember calculation. Samples SBB-1 and 46, SDC-1 and 2, SRV-4, XM-37 and HRV-147 all contain measurable K₂O. Samples SBB-1 and 46, SDC-1 and 2, XM-37 and HRV-147 all contain a normative calcium-Eskola ($\text{Ca}_{0.5}\text{AlSi}_2\text{O}_6$, CaEs) component, implying some M2 vacancy, with SBB-1 having almost 12 mol % of this endmember.

Infrared spectra at room temperature were obtained with a Nicolet 60SX Fourier Transform infrared spectrometer at a resolution of 2 cm^{-1} . Polished, self-supporting sample plates were mounted over pinhole apertures to define the viewing area through the sample. Viewing areas were 100–200 μm diameter. Pyroxene crystals were oriented using an X-ray precession camera, and (010) parallel facets were cut and polished. For each crystal the α - and γ -polarized spectra were measured in the X and Z principal vibration directions, respectively, in the

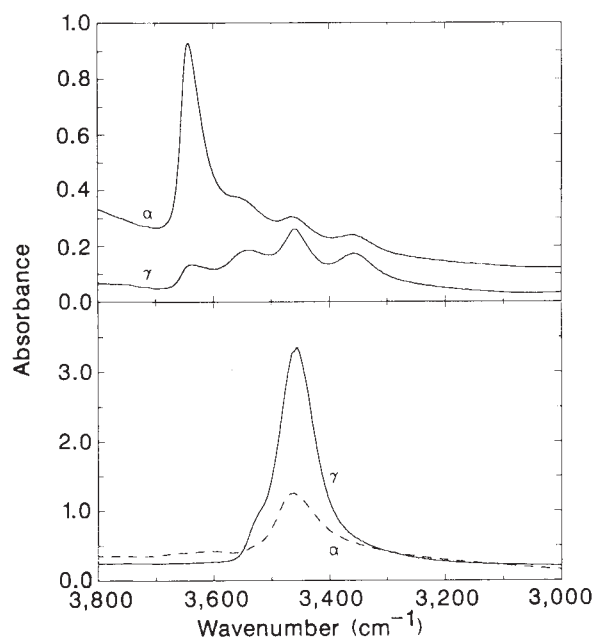


FIG. 1 Two representative infrared spectra in the range 3,000–3,800 cm^{-1} . a, Calcic sample (diopside) SRV-4 ($\text{Di}_{85}\text{Hd}_6\text{Ac}_5\text{CaTs}_3$), b, sodic sample (omphacite) SBB-1 ($\text{Jd}_{59}\text{Di}_{41}\text{Hd}_3\text{CaEs}_{12}$). The strong γ -polarized absorption band at 3,470 cm^{-1} correlates with the M2 vacancy concentration calculated from electron microprobe analysis.

region 3,000–3,800 cm^{-1} using a LiIO_3 crystal polarizer. (The X direction is approximately aligned with the crystallographic a -axis and Z is inclined $\sim 40^\circ$ from the c -axis.)

The spectra differed markedly for different compositions with a strong γ -polarized absorption band at 3,470 cm^{-1} appearing in jadeite-rich compositions and a strong α -polarized band at 3,620 cm^{-1} in the diopside-rich (Di-rich) compositions. Two representative spectra, one from a calcic sample SRV-4 of composition $\text{Di}_{85}\text{Hd}_6\text{Ac}_5\text{CaTs}_3$ (Hd is hedenbergite, Ac acmite, CaTs calcium-Tschermaks) and another from a sodic sample, $\text{Jd}_{59}\text{Di}_{20}\text{Hd}_8\text{CaEs}_{12}$, are plotted in Fig. 3. The relative intensities of the α absorption bands at 3,470 and 3,620 cm^{-1} , and the γ band at 3,470 cm^{-1} were calculated in absorbance units per mm and are listed in Table 2. From a calibration of infrared absorbance against OH content, we estimate the OH content of sample SBB-1 to be 1,840 p.p.m. by weight. The OH contents of the other samples are reported in Table 2.

Clinopyroxenes from mantle eclogites have complex compositions. The samples of this suite may contain more than 5 mol % of eight different endmembers. In addition to jadeite, diopside and hedenbergite, kosmochlor, acmite, Ca-Tschermaks and Ca-Eskola are important constituents. Also, six of the samples contain measurable potassium, of which five also contain a normative CaEs component calculated from the proportion of vacancies (4.00 minus total cations). McCormick⁸ has confirmed the presence of vacancies in M2 using X-ray crystal structure refinement and analytical transmission electron microscopy. She also reported the absence of multiple-chain defects in sample SBB-1 using high-resolution transmission electron microscopy.

The infrared spectra of these clinopyroxenes are also complex. In particular, the strong γ -polarized absorption feature at 3,470 cm^{-1} is very strong in sample SBB-1, whereas it is very weak relative to the α -polarized feature at 3,620 cm^{-1} in SRV-4 (Fig. 1). These differences are correlated with composition. Figure 2 is a plot of the intensity of the γ -3,470 feature against the cation deficiency per six oxygens. The 'negative' cation deficiencies of some samples may result from the effect of ferric iron on the vacancy calculation. If samples with positive cation deficiency also contain ferric iron, the calculated cation deficiency would be underestimated. The correlation in Fig. 2 indicates that the cation vacancy at M2 controls the incorporation of the hydroxyl responsible for the dominant 3,470- cm^{-1} band.

Despite the compositional complexities, no clinopyroxenes from these rocks have been reported with space groups other than $C2/c$, consistent with their high temperatures of equilibration. In this space group, the O2 site, bonded to one tetrahedral cation, one M1 site and one M2 site, is under-bonded according

TABLE 2 Absorbance per mm of mantle eclogite clinopyroxenes

Sample	α -polarized		γ -polarized	Total	OH (p.p.m.)*
	3,620 cm^{-1}	3,470 cm^{-1}	3,470 cm^{-1}		
SBB-1	0.068	0.844	3.040	3.95	1840
SBB-5H	0.530	0.040	0.154	0.72	335
SBB-7P	0.611	0.308	0.353	1.27	595
SBB-14	0.120	0.384	0.230	0.73	340
SBB-20	0.400	0.170	0.184	0.75	350
SBB-46	0.138	0.357	1.020	1.52	710
SBB-104	0.290	0.100	0.148	0.54	250
SDC-1	0.675	0.500	1.080	2.26	1055
SDC-2	0.400	0.280	0.865	1.55	725
SRV-4	0.693	0.093	0.207	0.99	460
XM-37	0.025	0.324	0.841	1.19	555
HRV-147	0.124	0.596	1.853	2.57	1200

* Concentrations in p.p.m. by weight are determined by assuming the summed OH band intensity is proportional to OH content. The reported⁷ OH content of HRV-147 was used as a standard.

to its Pauling bond strength, and has the lowest electrostatic site potential in the jadeite structure¹³. If we assume that the proton in these vacancy-bearing pyroxenes is associated with the O2 oxygen and that its bond is parallel to the γ -polarization direction, the O-H bond would extend roughly towards the M2 cation site. This is the same orientation Beran¹⁴ proposed for the OH group that produces the 3,600- cm^{-1} band in the diopside spectrum.

Experimental studies have demonstrated that vacancies exist in pyroxene at the pressures and temperatures at which eclogite is stable¹⁵. The correlation of hydroxyl content with vacancy concentration in the pyroxene suggests that both components existed in the pyroxene under mantle conditions. Clinopyroxenes are stable constituents of most model mantle compositions to pressures of 12 GPa, so hydroxyl-bearing pyroxenes may be important constituents of the mantle in the depth region of 30–400 km.

If the observed hydroxyl was indeed present in the mantle, then the finding of substantial hydroxyl concentrations associated with M2 cation vacancies is pertinent to understanding the geochemistry content of mantle processes. With an estimated maximum hydroxyl content of 1,840 p.p.m., these are the most hydrous pyroxenes yet observed, and indicate that sodic clinopyroxenes may be an important host for hydroxyl in the upper mantle. Electron microprobe analyses indicate that about one-third of our eclogite sample suite from Bellsbank and more than half of our suite from Roberts Victor contain a considerable concentration of M2 cation vacancies (up to $\sim 9\%$ in some samples). For the entire suite from both of the localities, the vacancy-bearing pyroxenes are 5–10 times as abundant as all other nominally hydrous phases (phlogopite and amphibole), so that the hydrous pyroxenes seem to contain well over half of the total hydrogen in the sample suite.

Compositions of pyroxenes reconstructed from exsolved garnet and kyanite indicate that precursor, near-solidus pyroxene may have contained up to 18% vacant M2 sites¹¹. This is supported by thermodynamic calculations¹⁵ based on experimental studies, which indicate that high temperatures favour the occurrence of high concentrations of vacancies in clinopyroxenes within the eclogite stability field. Extrapolating the curve in Fig. 2, we would estimate that the precursor pyroxene contained as much as 4,000 p.p.m. OH. Although the amphibole and ubiquitous phlogopite in the eclogite xenoliths are generally believed to derive from metasomatism by infiltrating kimberlitic fluids¹⁶, re-equilibration of pyroxenes during cooling from their crystallization temperature to their observed equilibration conditions ($\sim 1,050$ to $1,250^\circ\text{C}$ in the 3–5 GPa range)^{9,10} could evolve sufficient hydroxyl to account for much of that observed in the secondary hydrous phases in these rocks. Re-equilibration

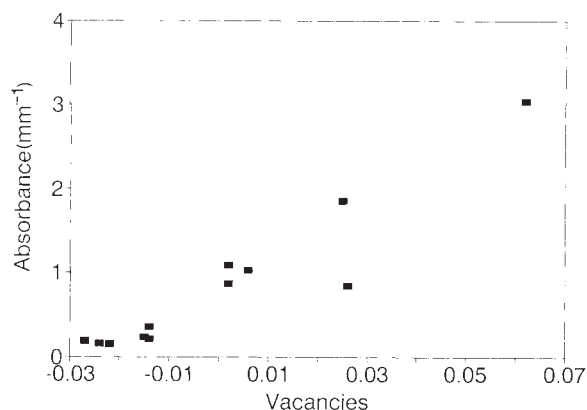


FIG. 2 Plot of absorbance normalized to 1 mm of the γ -polarized absorption peak at 3,470 cm^{-1} against the number of vacancies per 6.0 oxygens (calculated as 4.0 minus the total number of cations per six oxygens) in clinopyroxenes from mantle eclogites. In our calculations we have assumed that all iron is ferrous; a negative number may thus indicate the presence of some ferric iron.

could also result in a small amount of partial melt at depths of about 100 km. □

TABLE 1 Energy parameters used for simulation

	q	A (Å)	B (Å)	C (Å ³ kJ ^{1/2} mol ^{-1/2})
Mg	+1.565	1.0133	0.052	0
Si	+2.329	0.8436	0.040	0
O	-1.298	1.7600	0.150	54

These parameters are taken from ref. 10.

Received 10 January; accepted 13 May 1991.

- Beran, A. & Goetzinger, M. A. *Miner. Petrol.* **36**, 41–49 (1987).
- Navon, O., Hutcheon, I., Rossman, G. R. & Wasserburg, G. J. *Lunar planet. Sci.* **19**, 827–828 (1988).
- Rossman, G. R. & Smyth, J. R. *Am. Miner.* **75**, 775–780 (1990).
- Miller, G. H., Rossman, G. R. & Harlow, G. E. *Phys. Chem. Miner.* **14**, 461–472 (1987).
- Aines, R. D. & Rossman, G. R. *Geology* **12**, 720–723 (1984).
- Bell, D. R. & Rossman, G. R. *Trans. Am. geophys. Un.* **71**, 523 (1990).
- Skogby, H., Bell, D. R. & Rossman, G. R. *Am. Miner.* **75**, 764–774 (1990).
- McCormick, T. C. *Am. Miner.* **71**, 1434–1440 (1986).
- Smyth, J. R. & Caporuscio, F. A. in *Kimberlites II: The Mantle and Crust–Mantle Relationships* (ed. Kornprobst, J.) 121–132 (Elsevier, Amsterdam, 1984).
- Caporuscio, F. A. & Smyth, J. R. *Contrib. Miner. Petrol.* **105**, 550–561 (1990).
- Smyth, J. R., McCormick, T. C. & Caporuscio, F. A. *Kimberlites II: The Mantle and Crust–Mantle Relationships* (ed. Kornprobst, J.) 111–120 (Elsevier, Amsterdam, 1984).
- Hatton, C. J. thesis, Univ. Cape Town (1978).
- Smyth, J. R. *Geochim. cosmochim. Acta* **53**, 1101–1110 (1989).
- Beran, A. *Tschermaks Min. Petrol. Mitt.* **23**, 79–85 (1976).
- Gasparik, T. *Am. Miner.* **71**, 686–693 (1986).
- Taylor, L. A. & Neal, C. R. *J. Geol.* **97**, 551–567 (1989).

ACKNOWLEDGEMENTS. We thank J. J. Gurney and J. B. Hawthorne for samples and T. McCormick for helpful discussions. This study was funded in part by the NSF.

Simulation of the pre-melting behaviour of MgSiO₃ perovskite at high pressures and temperatures

Masanori Matsui* & Geoffrey D. Price

Department of Geological Sciences, University College London, Gower Street, London WC1E 6BT, UK

MAGNESIUM-rich silicate perovskite is thought to be the dominant mineral phase in the Earth's lower mantle. The behaviour of MgSiO₃ perovskite at high temperatures and pressures is therefore important for a wide range of geophysical problems, including the chemical and thermal evolution of the Earth, mantle convection, the thermal gradient in the mantle and the secular variations of the Earth's magnetic field. Experimental investigations at lower-mantle conditions are, however, difficult. We have performed computer simulations of MgSiO₃ perovskite under typical lower-mantle pressures and temperatures using the constant-temperature and constant-pressure molecular dynamics (MD) method. At pressures above 10 GPa, our simulations suggest that orthorhombic MgSiO₃ perovskite undergoes a temperature-induced phase transformation to a cubic (or pseudo-cubic) phase before melting, and that the cubic phase is a solid electrolyte. The MD method tends to overestimate the temperature of melting and related phenomena, but should provide a reliable qualitative description of the ionic-conductivity behaviour. Quantitative determination of the structure and ionic conductivity of MgSiO₃ perovskite at lower-mantle conditions must, however, await improvements in both experimental and simulation techniques.

There is controversy concerning the electrical conductivity of (Mg, Fe)SiO₃ perovskite under lower-mantle temperatures and pressures. Li and Jeanloz^{1,2} have measured the electrical conductivity of (Mg, Fe)SiO₃ perovskite at typical lower-mantle temperatures and pressures, and concluded that (Mg, Fe)SiO₃ perovskite is an insulator throughout the lower mantle. In contrast, Poirier and Peyronneau³ have inferred, from the extrapolation of conductivity data measured at pressures of 40 GPa and temperatures up to 673 K, that (Mg, Fe)SiO₃ perovskite would have an electrical conductivity of 1 S m⁻¹ at a depth of

1,100 km in the lower mantle and of 50–100 S m⁻¹ at the core-mantle boundary. The conclusions of Poirier and Peyronneau³ have been recently reinforced by Wood and Nell⁴, from a study of the conductivity of magnesiowüstite. In addition, recent MD studies^{5,6} have indicated that MgSiO₃ perovskite might be an ionic conductor under mantle conditions. These simulations, however, were carried out under constant-volume conditions, which are unsatisfactory for simulating materials that are undergoing a phase transformation. Here we present the results of a constant-pressure, constant-temperature MD simulation of phase transformations in MgSiO₃ perovskite. We used a combination of the constant-temperature MD method proposed by Nosé⁷ and the constant-pressure MD technique developed by Parrinello and Rahman⁸. Quantum effects⁹ on the calculated properties are unimportant at the high temperatures considered in this present study.

Our simulations were based on a cell, composed of 27(3a × 3b × 3c) orthorhombic perovskite unit cells, which contained 540(108 Mg, 108 Si, and 324 O) atoms. The equations of motion were solved by a fifth-order predictor–corrector algorithm with a time increment of 1.0 fs. All ions interacted through a rigid-ion two-body potential¹⁰, $\Phi_{ij}(r) = q_i q_j r_{ij}^{-1} - C_i C_j r_{ij}^{-6} + f(B_i + B_j) \exp[(A_i + A_j - r_{ij})/(B_i + B_j)]$, where r is the distance between atoms i and j , and f is a force constant of 4.184 kJ Å⁻¹ mol⁻¹. The potential parameters used in our simulations are listed in Table 1. Interatomic separations during the simulation were monitored to ensure that no ion sampled the artificially attractive portion of the Buckingham potential, known to occur for small values of r_{ij} . The potential used has been found to reproduce not only the observed structure, volume compressibility and thermal expansivity of MgSiO₃ perovskite¹⁰, but also the observed structures, densities and enthalpy trends of all the known MgSiO₃ polymorphs: orthoenstatite, clinoenstatite, protoenstatite, ilmenite, garnet and perovskite itself.

We simulated the temperature-induced phase transformation and melting of MgSiO₃ perovskite over the pressure range 0–150 GPa. Under each fixed pressure condition, the temperature was increased gradually near the transformation temperature, in steps of 100 or 200 K, to minimize kinetic effects. We found no evidence that the observed phase transformations were being significantly over-driven, and indeed the transformation from insulator to solid electrolyte described below was reproducible with a hysteresis of only 100–200 K. In each MD run, at fixed temperature and pressure, the simulation was performed for 5,000 time steps (5 ps) to equilibrate the system, followed by a further 5,000 time steps from which the statistical averages of the structural properties were obtained. During the course of each MD run, the mean square displacement at time t was calculated for each species, and the diffusion coefficient was obtained from the slope¹¹ of the mean square displacement with respect to time.

Figure 1 shows the simulated temperature dependence of the unit cell parameters and the molar volume of MgSiO₃ perovskite at a pressure of $P = 30$ GPa (chosen as a typical example); in the ideal cubic perovskite structure, the unit cell parameters in the orthorhombic setting satisfy the relation $a = b = c/\sqrt{2}$. As can be seen in Fig. 1, the three cell parameters a , b and $c/\sqrt{2}$ become progressively closer to each other with rising temperature and become identical, within the limits imposed by the

* Permanent address: Chemical Laboratory, Kanazawa Medical University, Ishikawa 920-02, Japan.

fluctuations of the system, at about 4,800 K. At lower temperatures, equilibrations were quite stable, whereas the results above 4,200 K fluctuated considerably over the timescale used for simulation. In addition, on transforming from the orthorhombic to the cubic (or pseudo-cubic) phase, MgSiO_3 perovskite became a solid electrolyte, in which the oxygen ions were disordered and highly mobile, but the cations remained at their crystallographic sites. The simulated diffusion coefficient of the oxygen ions at 5,000 K and 30 GPa was $1 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$, which is comparable with the diffusion coefficients of ions in molten salts. Similar high conductivities were found in the two previous constant volume simulations^{5,6} of MgSiO_3 perovskite. These simulations were based on potentials that are markedly different from the one used here, suggesting that the prediction of the superionic behaviour of MgSiO_3 perovskite is not highly dependent on the details of the potential used.

In our model, both the temperature of the solid electrolyte transition and the melting temperature are predicted to increase with increasing pressure. The orthorhombic-to-cubic phase transformation and the associated solid electrolyte behaviour was only found to occur above 10 GPa. At pressures less than this the orthorhombic phase was predicted to melt before any structural transformation occurred. Unlike Stixrude and Bukowski¹², we predict no maximum in the liquidus curve. Indeed, at no pressure did the density of the simulated melt exceed that of the solid from which it was derived. An analysis of our calculated radial distribution functions shows that silicon coordination in the melt increases gradually from four to six

with increasing pressure, but there is no evidence for the adoption of a coordination significantly higher than octahedral.

Temperature-induced phase transformations from orthorhombic to cubic, through tetragonal, have been predicted for MgSiO_3 perovskite by lattice dynamics¹³, observed experimentally in NaMgF_3 perovskite¹⁴ and simulated by constant-pressure MD techniques for RbCaF_3 ¹⁵ and KNNF_3 ¹⁶ perovskites. Interestingly, on the basis of electrical conductivity measurements at high temperatures, the cubic phase of NaMgF_3 perovskite was also reported¹⁷ to be a solid electrolyte. KZnF_3 perovskite, with cubic symmetry, was also reported¹⁸ to be a solid electrolyte, on the basis of viscosity and electrical conductivity measurements at high temperatures. From these observations and from our simulation, it seems that the development of superionic conductivity is closely linked to the symmetry of the perovskite phase. The transformation to cubic symmetry may be caused by the onset of ionic mobility, or alternatively it may be that cubic symmetry is a necessary (but perhaps not sufficient) condition for solid electrolyte behaviour.

The predicted melting point of MgSiO_3 perovskite at 25 GPa (4,800 K) is substantially higher than the reported experimental value¹⁹ ($3,000 \pm 300$ K) for $\text{Mg}_{0.9}\text{Fe}_{0.1}\text{SiO}_3$ perovskite. It is also higher than the previous estimations, based on the Kraut-Kennedy or Simon equation²⁰, the Lindemann melting law²¹, or the estimated difference in Helmholtz free energy between the solid and liquid phases¹². But as we have said, the potential used succeeds in reproducing the structural and energetic properties of all the MgSiO_3 polymorphs fairly accurately. We therefore believe that the high predicted melting temperature of perovskite (and of the other MgSiO_3 polymorphs studied) is not due to an inherently poor description of the interatomic short-range interactions, but may be because we have ignored surface effects (which should certainly lower the simulated melting temperature), or because the rigid-ion potentials used in this simulation overestimate the enthalpy of the MgSiO_3 melt. The ions in melts are mobile and highly disordered when compared with the crystal. Catlow *et al.*²² have pointed out that ionic diffusion and the defect-related properties of a system are not generally well reproduced by the rigid-ion model, and that some explicit representation of ionic polarizability (for example by the shell model) is necessary to simulate atomic relaxation around a defect and to give an adequate estimate of the defect energy. Indeed, we have previously successfully used polarizable oxygen shell-models to simulate the defect behaviour of MgSiO_3 perovskite^{23,24}. But including a shell model in a MD simulation is computationally very expensive and is not a proven method.²⁵ It was impossible therefore to use our more-sophisticated potentials in these MD simulation studies. Neither is it possible to use the quasi-harmonic lattice dynamics technique, which can support more sophisticated potentials, as this approach is valid only at relatively low temperatures, and it is incapable of simulating the highly anharmonic behaviour associated with melting. Finally, the size of the box used in our MD simulations may affect the predicted melting temperatures, although no significant change in predicted behaviour was observed when the number of particles in the simulation was doubled.

Our simulations show that orthorhombic MgSiO_3 perovskite transforms to a solid electrolyte phase with a cubic unit cell at lower-mantle pressures, although the predicted transformation temperatures are high when compared with the probable range of lower-mantle temperatures. But as our rigid-ion model almost certainly overestimates the defect energy of the system, and as atomic transport in solids is determined by defect migration, it seems likely that the predicted transformation temperatures is overestimated to the same degree as the predicted perovskite melting temperature. With the present limitations of the MD technique, therefore, it is not possible to model the structural state of MgSiO_3 perovskite unambiguously under mantle conditions. We suggest, however, that whether MgSiO_3 perovskite shows solid electrolyte behaviour in the lower mantle will

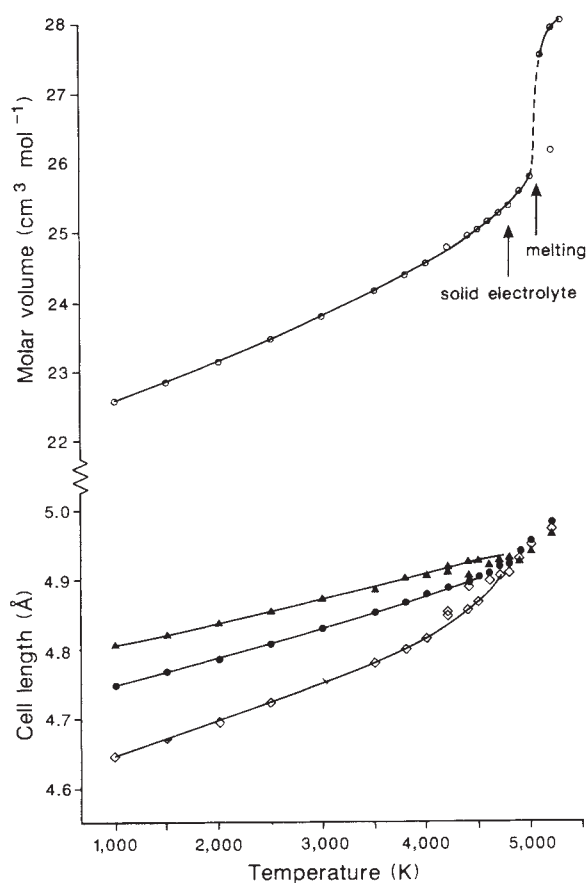


FIG. 1 Temperature dependence of the calculated cell parameters (lower part) and molar volume (upper part) of MgSiO_3 perovskite at 30 GPa, showing the orthorhombic-to-cubic phase transformation at 4,800 K and melting at 5,100 K. The simulated points in the cubic phase at 5,200 K show a superheating effect. Pressure $P=30$ GPa. Cell parameters $\diamond$, a ; $\bullet$, $c/\sqrt{2}$; $\blacktriangle$, b .

depend on whether it exists in the orthorhombic (insulating) or in the cubic (conducting) form. □

Received 11 March; accepted 2 May 1991.

1. Li, X. & Jeanloz, R. *Geophys. Res. Lett.* **14**, 1075–1078 (1987).
2. Li, X. & Jeanloz, R. *J. geophys. Res.* **95**, 5067–5078 (1990).
3. Poirier, J. P. & Peyronneau, J. *Nature* **342**, 537–539 (1989).
4. Wood, B. J. & Nell, J. *Nature* **351**, 309–311 (1991).
5. Kapusta, B. & Guillope, M. *Phil. Mag.* **58**, 809–816 (1988).
6. Wall, A. & Price, G. D. *Phys. Earth planet. Inter.* **58**, 192–204 (1989).
7. Nose, S. *J. chem. Phys.* **81**, 511–519 (1984).
8. Parrinello, M. & Rahman, A. *J. appl. Phys.* **52**, 7182–7190 (1981).
9. Matsui, M. *J. chem. Phys.* **91**, 489–494 (1989).
10. Matsui, M. *Phys. Chem. Miner.* **16**, 234–238 (1988).
11. Gillan, M. J. & Dixon, M. *J. Phys. C* **13**, 1901–1917 (1980).
12. Stixrude, L. & Bukowski, M. S. T. *J. geophys. Res.* **95**, 19311–19325 (1990).
13. Wolf, G. H. & Bukowski, M. S. T. in *High Pressure Research in Mineral Physics* (eds Manghni, M. H. & Syono, Y.) 313–331 (Terra Scientific, Tokyo/Am. Geophys. Un., Washington DC, 1987).
14. Chao, E. C. T., Evans, H. T., Skinner, B. J. & Milton, C. *Am. Miner.* **46**, 379–393 (1961).
15. Nosé, S. & Klein, M. L. *J. chem. Phys.* **90**, 5005–5010 (1989).
16. Lewis, L. J. & Lépine, Y. *Phys. Rev. B* **40**, 3319–3322 (1989).
17. O'Keeffe, M. & Bovin, J.-O. *Science* **206**, 599–600, (1979).
18. Poirier, J. P., Peyronneau, J., Gesland, J. Y. & Brebec, G. *Phys. Earth planet. Inter.* **32**, 273–287 (1983).
19. Heinz, D. L. & Jeanloz, R. *J. geophys. Res.* **92**, 11437–11444 (1987).
20. Ohtani, E. *Phys. Earth planet. Inter.* **33**, 12–25 (1983).
21. Poirier, J. P. *Phys. Earth planet. Inter.* **54**, 364–369 (1989).
22. Catlow, C. R. A., Dixon, M. & Mackrodt, W. C. in *Computer Simulation of Solids* (eds Catlow, C. R. A. & Mackrodt, W. C.) 130–161 (Springer, Berlin, 1982).
23. Wall, A. & Price, G. D. in *Perovskites* (eds Navrotsky, A. & Weidner, D. J.) 45–53 (Am. Geophys. Un., Washington DC, 1989).
24. Price, G. D., Wall, A. & Parker, S. C. *Phil. Trans. R. Soc. Lond. A* **328**, 391–407 (1989).
25. Saboungi, M.-L., Rahman, A., Halley, J. W. & Blander, M. *J. chem. Phys.* **88**, 5818–5823 (1988).

ACKNOWLEDGEMENTS. We thank R. Angel and J.-P. Poirier for many interesting discussions, and J. Kubicki for helpful comments on the manuscript. We thank the Royal Society for supporting M. M. visit to the UK and the NERC for the provision of supercomputing resources.

Thermoluminescence dating of the late Neanderthal remains from Saint-Césaire

N. Mercier, H. Valladas, J.-L. Joron*, J.-L. Reyss, F. Lévêque† & B. Vandermeersch‡

Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, Avenue de la Terrasse, 91198 Gif-sur-Yvette Cedex, France

* Groupe des Sciences de la Terre, Laboratoire Pierre Süe, CEN, Saclay, 91191 Gif-sur-Yvette, France

† Direction des Antiquités de la Région Poitou-Charentes, Hôtel de Rochefort, 102 Grand'Rue, BP 553, 86020 Poitiers, France

‡ Laboratoire d'Anthropologie, Université de Bordeaux 1, UA 376 du CNRS, 33405 Talence, France

ANATOMICALLY modern humans have long been thought to have been responsible for the Aurignacian and Châtelperronian industries of the early Upper Palaeolithic of Western Europe, whereas the Middle Palaeolithic Mousterian industry has been attributed to Neanderthals. The presence of both Middle and Upper Palaeolithic strata at Saint-Césaire in France offers an excellent opportunity for studying the cultural transition between the two. Saint-Césaire is the only Châtelperronian site that has yielded really diagnostic hominid fossils, and the discovery there of Neanderthal remains¹ alongside Châtelperronian tools cast doubt on the exclusive association between industries and taxon. We report thermoluminescence dates for 20 burnt flints from the site. Those found near the Neanderthal remains were dated at $36,300 \pm 2,700$ years BP (before present), making this specimen the youngest Neanderthal dated so far. This date places the stratum close in age to several French^{2,3} but much younger than some Spanish^{4,5} Aurignacian sites believed to have been occupied by modern humans. The possibility of contact between the West European Neanderthals and the intrusive modern humans who replaced them cannot therefore be excluded.

The biological and archaeological aspects of the transition from Middle to Upper Palaeolithic remain a controversial sub-

ject^{6–8}. The European Neanderthals were for a long time associated with Middle Palaeolithic Mousterian industries, whereas anatomically modern humans were thought to be responsible for the Early Upper Palaeolithic Aurignacian and Châtelperronian lithic industries, despite the affinities the latter showed with 'Mousterian' or 'Acheulian' Tradition^{9,10}. The correlation between the specified hominids and industries has been complicated by the presence of a Neanderthal of a fairly classical type, judging by the morphology of the skull and postcrania, in a Châtelperronian stratum at Saint-Césaire¹¹. The discovery has confirmed the growing suspicion, first aroused by the presence of possibly Neanderthal teeth at Arcy-sur-Cure^{12–14} that early Upper Palaeolithic Châtelperronian industries in French and Spanish sites¹⁵ were made by late Neanderthals, hence the importance of dating this crucial phase in the transition from Middle to Upper Palaeolithic.

The collapsed rock shelter of 'La Roche à Pierrot', near the village of Saint-Césaire, was excavated by Lévêque (1976–1987)¹⁶. It yielded Mousterian, Châtelperronian and Aurignacian lithic remains. The 2.5 m-deep deposit, subdivided into 17 levels, contained two distinct units. The first, 'grey unit (Eg)', consisted of eight mostly sandy Mousterian layers. The Mousterian of Acheulian Tradition was well represented in layers 16–13, and Denticulate Mousterian in layers 12–10.

Clay predominated in the second, 'yellow unit (Ej)', which contained seven Upper Palaeolithic layers, among them two Châtelperronian (9–8). Covering these was a layer (7) of sterile soil, over which four Aurignacian layers were subsequently deposited. The latter consisted of one proto-Aurignacian (6), one Early Aurignacian (5) and two Late Aurignacian (4, 3). None of the excavated bone remains contained enough collagen to give a reliable radiocarbon date (M. Fontugne, personal communication) and no charcoal has been recovered from the site. Among the flints collected in the course of the excavation, about 50 showed evidence of burning but only 20 were satisfactory for thermoluminescence dating. Two came from the proto-Aurignacian layer 6, six from the Châtelperronian layer 8, in which the Neanderthal bones were found, and the rest from the underlying Mousterian strata.

The sample preparation and the experimental procedure have been described previously^{17–19}. The external radiation dose was calculated from data recorded by 40 dosimeters (CaSO₄/Dy) planted at various levels. The γ dose decreased with depth in close agreement with the mineralogical composition of the sediments. The doses received by flints from the different layers are shown in Table 1, where all relevant experimental data are tabulated. The internal radiation dose, computed from the concentrations of radioelements (determined by neutron activation analysis), accounts for 40–63% of the total annual dose. This reduces the impact of variations in external dose (which range from ± 15 to $\pm 25\%$) on the calculated age.

The dates of individual flints with the statistical errors are shown as a function of depth in Fig. 1, accompanied by the weighted average and the total (statistical + systematic) error for each layer. The averages, ranging from 42,000 to 32,000 yr BP, fit the stratigraphic evidence. The Denticulate Mousterian layers (10–12) cluster about 40,000 yr BP, suggesting that they were deposited over a short time interval. The date of $36,300 \pm 2,700$ yr BP for the layer containing the Neanderthal remains is notable as no other Neanderthal bones dated so far are as late. The gap of over 2,000 years separating the uppermost Mousterian from the dated Châtelperronian can be explained in part by a stratigraphic discontinuity^{15,20} apparently caused by erosion of the layers overlying the surviving Mousterian. The proto-Aurignacian layer is $32,100 \pm 3,000$ years old.

Although thermoluminescence gives the date for the Châtelperronian stratum, which enclosed the Neanderthal, as older than the radiocarbon dates reported for the same industry at Arcy-sur-Cure ($33,860 \pm 250$ yr BP)²¹ and Les Cottés ($33,300 \pm 500$ yr BP)², the difference may not be as important in view of

TABLE 1 Thermoluminescence results and radioactivity data for the flints from Saint-Césaire

Sample	U*	Th*	K*	S α †	D α ‡	D β §	Internal dose	External dose¶	Annual dose	Palaeodose (Gray)	Age (kyr BP)
	(p.p.m.)						($\mu\text{Gy yr}^{-1}$)				
Layer 6-Ejo sup											
96	0.530	0.491	666	1.99	257 $\pm$ 8	146	408 $\pm$ 25	610	1,020 $\pm$ 100	31.23 $\pm$ 1.2	30.8 $\pm$ 3.3
60	0.541	0.237	430	1.45	170 $\pm$ 10	121	296 $\pm$ 20	610	900 $\pm$ 100	30.70 $\pm$ 1.03	34.0 $\pm$ 3.9
Layer 8-Ejop sup											
95	0.687	0.466	1,599	1.64	259 $\pm$ 33	244	510 $\pm$ 45	375	890 $\pm$ 110	32.39 $\pm$ 1.73	36.6 $\pm$ 5.0
48	0.836	0.197	178	1.29	222 $\pm$ 13	142	374 $\pm$ 26	395	770 $\pm$ 100	29.35 $\pm$ 1.04	38.2 $\pm$ 5.3
103	0.485	0.106	378	1.84	183 $\pm$ 6	104	291 $\pm$ 18	407	700 $\pm$ 100	23.48 $\pm$ 1.33	33.7 $\pm$ 5.4
53	0.614	0.029	240	2.71	327 $\pm$ 16	110	441 $\pm$ 31	402	850 $\pm$ 100	30.82 $\pm$ 1.38	36.6 $\pm$ 4.9
54	0.510	0.069	430	2.48	254 $\pm$ 13	111	370 $\pm$ 26	402	770 $\pm$ 100	28.86 $\pm$ 0.93	37.4 $\pm$ 5.2
82	0.719	0.397	484	1.85	297 $\pm$ 19	156	460 $\pm$ 33	418	880 $\pm$ 110	31.30 $\pm$ 1.08	35.6 $\pm$ 4.6
Layer 10-Egpf											
75	0.529	0.188	274	1.49	167 $\pm$ 11	105	277 $\pm$ 20	351	630 $\pm$ 60	27.06 $\pm$ 1.26	43.1 $\pm$ 4.8
107	0.946	0.260	732	1.77	350 $\pm$ 18	205	565 $\pm$ 38	352	920 $\pm$ 70	36.20 $\pm$ 1.95	39.5 $\pm$ 3.9
106	0.400	0.415	1,206	1.69	168 $\pm$ 3	169	341 $\pm$ 20	359	700 $\pm$ 60	27.29 $\pm$ 1.75	39.0 $\pm$ 4.5
83	0.616	0.270	292	1.74	233 $\pm$ 7	121	359 $\pm$ 22	281	640 $\pm$ 60	21.40 $\pm$ 1.38	33.5 $\pm$ 3.8
66	0.461	0.120	440	1.79	171 $\pm$ 12	106	282 $\pm$ 21	315	600 $\pm$ 50	25.35 $\pm$ 1.18	42.4 $\pm$ 4.4
105	1.051	0.212	864	1.48	318 $\pm$ 15	230	555 $\pm$ 36	318	870 $\pm$ 60	37.92 $\pm$ 1.25	43.4 $\pm$ 3.4
111	0.652	0.289	780	1.93	274 $\pm$ 19	167	447 $\pm$ 33	315	760 $\pm$ 60	29.64 $\pm$ 1.02	38.9 $\pm$ 3.4
112	0.603	0.342	842	1.30	175 $\pm$ 14	166	348 $\pm$ 25	336	680 $\pm$ 70	31.19 $\pm$ 1.39	45.6 $\pm$ 4.8
25	0.472	0.389	610	1.84	206 $\pm$ 11	130	344 $\pm$ 23	310	650 $\pm$ 60	30.82 $\pm$ 1.56	47.1 $\pm$ 4.9
Layer 11-Egp											
84	0.746	0.384	472	1.14	188 $\pm$ 13	158	353 $\pm$ 24	257	610 $\pm$ 60	24.22 $\pm$ 0.91	39.7 $\pm$ 3.9
78	0.452	0.337	364	1.67	176 $\pm$ 8	105	286 $\pm$ 19	284	570 $\pm$ 50	20.96 $\pm$ 0.74	36.8 $\pm$ 3.7
Layer 12-Egf											
86	0.629	0.272	320	1.49	203 $\pm$ 13	125	340 $\pm$ 24	251	590 $\pm$ 60	25.06 $\pm$ 1.00	42.4 $\pm$ 4.3

* The values of U, Th and K were measured by neutron activation analysis at the Institut Pierre S   (CEN, Saclay) and each has an error of $\pm 6\%$ ²⁵.

† β -dose equivalent to the flux of $1\alpha\text{ cm}^{-2}$, as deduced from fine grains measurements²⁶. The error on S α ranges from $\pm 2\%$ to $\pm 12\%$.

‡ β -dose equivalent to the α dose in flint calculated using the specific annual flux deduced from ref. 27.

§ The β -dose in flints was calculated using the specific values given in ref. 28. The γ internal dose is less than 0.01 mGy yr^{-1} .

¶ The external dose includes a cosmic contribution ranging from 0.10 to 0.13 mGy yr^{-1} , calculated using the data given in ref. 29.

The uranium series radionuclides are essentially at equilibrium in the sediment: the activities of ²³⁴Th, ²²⁶Ra, ²¹⁴Bi and ²¹⁰Pb measured by γ spectrometry are 1.13 ± 0.09 , 0.99 ± 0.08 , 0.90 ± 0.05 and $1.05 \pm 0.14\text{ d.p.m. g}^{-1}$, respectively.

recent evidence^{22,23} that ¹⁴C dates from this period might be in error by as much as 10% too young. This same evidence would make the surviving remains of Neanderthal of Saint-C  saire a contemporary of the occupants of several radiocarbon-dated Aurignacian sites, such as L'Abri Pataud ($34,250 \pm 675\text{ yr BP}$)² or Le Flageolet ($33,800 \pm 1,800\text{ yr BP}$)³, both currently believed to have been occupied by modern humans²⁴. Moreover, even if the recently reported radiocarbon dates for the Aurignacian levels at El Castillo ($38,700 \pm 1,900\text{ yr BP}$)⁴ and L'Arbreda

($38,500 \pm 1,000\text{ yr BP}$)⁵ caves in Spain are not adjusted, the thermoluminescence date of the Saint-C  saire Neanderthal suggests that members of this population might have survived several millennia after the arrival of people responsible for the Aurignacian industries of Western Europe.

Because, despite its late date, the Saint-C  saire skull has classical Neanderthal features¹¹ with no evidence of evolution in the direction of modern humans, the idea that the former became extinct by evolving into the latter becomes chrono-

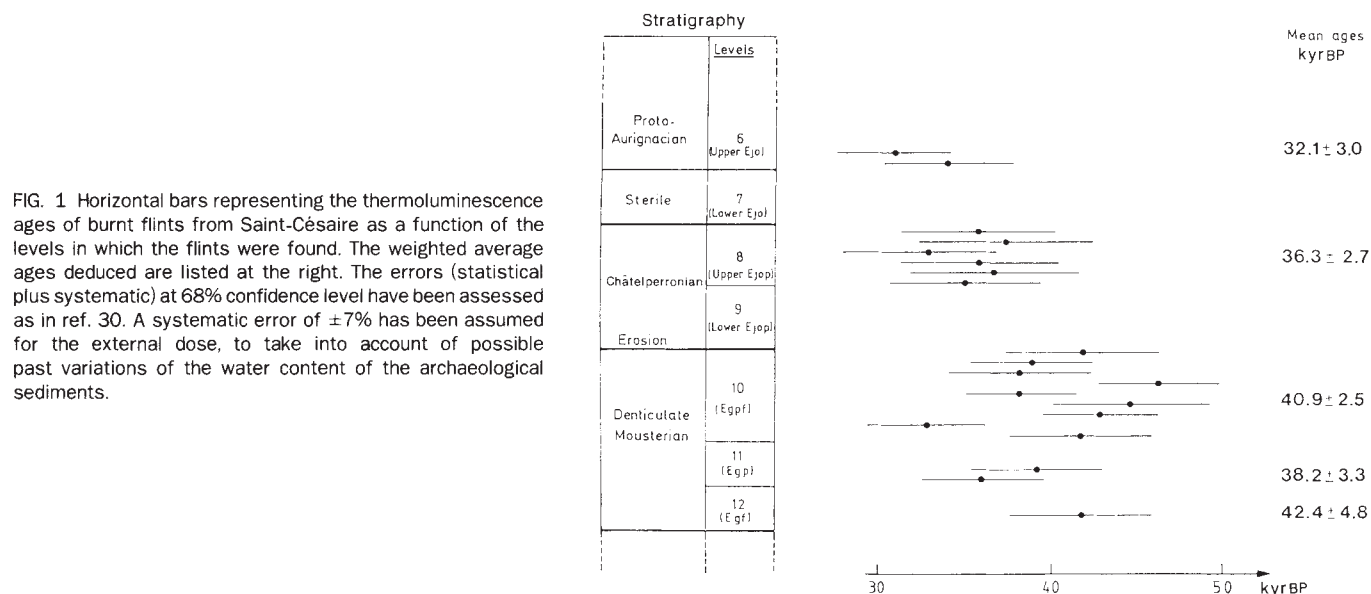


FIG. 1 Horizontal bars representing the thermoluminescence ages of burnt flints from Saint-C  saire as a function of the levels in which the flints were found. The weighted average ages deduced are listed at the right. The errors (statistical plus systematic) at 68% confidence level have been assessed as in ref. 30. A systematic error of $\pm 7\%$ has been assumed for the external dose, to take into account of possible past variations of the water content of the archaeological sediments.

logically untenable, at least in Western Europe. But the Saint-Césaire dates still leave unanswered the question: do the Châtelperronian industries represent the last Neanderthal innovations in a long Mousterian tradition, and should they thus be considered as defining the end of the Middle Palaeolithic; or were they created through contact with the newly arrived modern humans and thus do they mark the beginning of the Upper Palaeolithic³¹?

Note added in proof. In the latter case, considering the temporal overlap between the two populations, as indicated by our dates, the Châtelperronian industry could have arisen through acculturation of the Neanderthals in contact with modern humans. □

Received 26 March; accepted 13 May 1991.

1. Lévêque, F. & Vandermeersch, B. *C. R. Acad. Sci., Paris* **291**, 187–189 (1980).
2. Vogel, J. C. & Waterbolk, H. T. *Radiocarbon* **9**, 107–156 (1967).
3. Gowlett, J. A. J., Hedges, R. E. M., Law, I. A. & Perry, C. *Archaeometry* **28**, 206–222 (1986).
4. Cabrera-Valdés, V. & Bischoff, J. L. *J. arch. Sci.* **16**, 577–584 (1989).
5. Bischoff, J. L., Soler, N., Maroto, J. & Julia, R. *J. arch. Sci.* **16**, 563–576 (1989).
6. Mellars, P. *Curr. Anthropol.* **30**, 349–384 (1989).
7. Demars, P. Y. & Hublin, J.-J. in *l'Homme de Néandertal, 3: l'Extinction* (ed. Otte, M.) 23–37 (Eraul 34, Liège, 1989).
8. Mellars, P. & Stringer, C. (eds) *The Human Revolution* (Edinburgh University Press, Edinburgh, 1989).
9. Leroi-Gourhan, A. *Bull. Soc. Mériod. Spéleol. Préhist.* 6–9, (1963).
10. Bordes, F. *Le Paléolithique dans le Monde, l'Univers des Connaissances*, (Hachette, Paris, 1968).
11. Vandermeersch, B., *Bull. Mém. Soc. d'Anthrop. de Paris* **1**, 191–196 (1984).
12. Leroi-Gourhan, A. *Préhistoire de l'Art Occidental* (ed. Mazenod, L.) (Paris, 1965).
13. ApSimon, A. M. *Nature* **287**, 272–273 (1980).
14. Wolpoff, M. H.; ApSimon, A.; Stringer, C. B.; Kruszynski, R. G. & Jacobi, R. M. *Nature* **289**, 823–824 (1981).
15. Lévêque, F. & Miskovsky, J. C. *L'Anthropologie* **87**, 369–391 (1983).
16. Lévêque, F. in *l'Homme de Néandertal, 3: l'extinction* (ed. Otte, M.) 99–108 (Eraul 34, Liège, 1989).
17. Aitken, M. J., *Thermoluminescence Dating* (Academic, New York, 1985).
18. Valladas, H. et al. *Nature* **322**, 452–454 (1986).
19. Valladas, H. & Valladas, G. *Archaeometry* **27**, 214–220 (1987).
20. Leroi-Gourhan, A. *Bull. Soc. Préhist. France* **81**, 196–198 (1984).
21. Girard, M., Miskovsky, J. C., & Evin, J. *Mém. Mus. Préhist. d'Île de France* **3**, 295–303 (1990).
22. Bard, E. et al. *Nature* **345**, 405–409 (1990).
23. Bell, W. T., *Archaeometry* **33**, 43–50 (1991).
24. Gambier, D. in *l'Homme de Néandertal, 3: l'Extinction* (ed. Otte, M.) 67–84 (Eraul 34, Liège, 1989).
25. Chayla, B., Jaffrezic, H. & Joron, J.-L. *C. R. hebdom. Acad. Sci., Paris* **277**, 273–275 (1973).
26. Huxtable, J. & Jacobi, R. M. *Archaeometry* **24**, 164–169 (1982).
27. Valladas, H. & Valladas, G. *Pact* **6**, 171–178 (1982).
28. Bell, W. T., *Archaeometry* **21**, 243–245 (1979).
29. Prescott, J. R. & Stephan, L. G. *Pact* **6**, 17–25 (1982).
30. Aitken, M. J. *Archaeometry* **18**, 233–238 (1976).
31. Vandermeersch, B., *Mém. Mus. Préhist. d'Île de France* **3**, 25–27 (1990).

Mortality rates and population density of tsetse flies correlated with satellite imagery

D. J. Rogers & S. E. Randolph

Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

TSETSE flies are a major constraint on animal production in about 10 million km² of Africa through their transmission of animal trypanosomiasis¹. Up to 25 million people are at risk from human trypanosomiasis, or sleeping sickness². Tsetse research has been concentrated on the factors that control the distribution and abundance of these vectors and the means by which their numbers can be reduced³. Eradication successes in some countries are insignificant compared with the continental scale of the problem and the long-term reduction in the area infested by tsetse has been negligible⁴. We report here that the mortality rates of tsetse from sites in both West and East Africa, the size of male and female tsetse (related to the mortality rate of the parental female population) along a north–south transect in West Africa, and the abundance of two species of tsetse over the northern half of Côte d'Ivoire, are significantly correlated with data from meteorological satellites. This information could be used to predict both the mortality rate and the abundance (key determinants of disease transmission potential) of tsetse over very large areas of the continent and to

produce maps of high risk areas of disease transmission for the African trypanosomiasis and, by implication, for many other vector-borne diseases.

Tsetse flies, *Glossina* spp., are sensitive to temperature, rainfall and saturation deficit. The monthly mortality rates of tsetse, estimated by applying Moran curve techniques⁵ to data derived from standardized fly-round⁶ or trap-catches of flies, combined with the climogram of the study site, have previously been used to predict the distribution of the species over much of Africa⁷. Figure 1a shows the relationship between the bimonthly (that is, the average of the present and previous months') density-independent mortality (expressed as a *k*-value⁸) of *G. morsitans submorsitans*, a savannah species of tsetse from Nigeria⁴, and the saturation deficit of the previous month. The fertility rate of tsetse, expressed on the same scale as the mortality rates

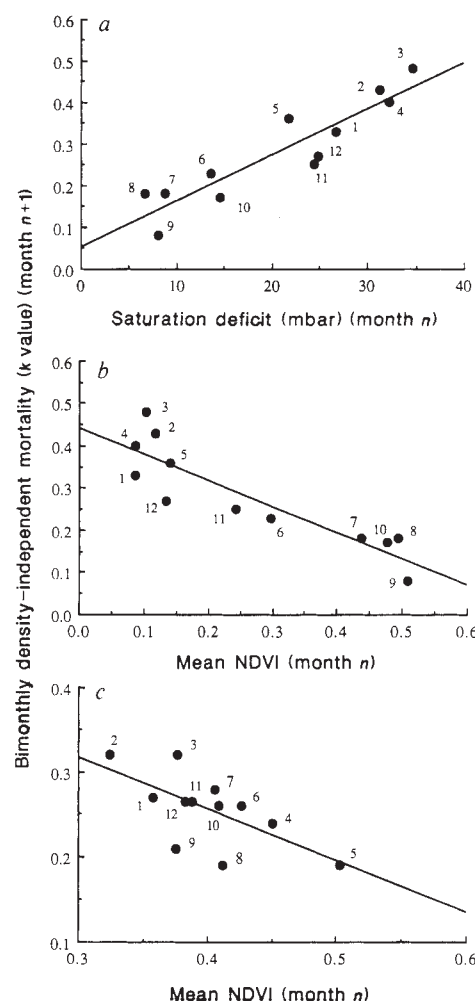


FIG. 1 Relationships between tsetse fly mortality rates and ground-based or satellite-derived estimates of environmental conditions. The relationship between the bimonthly density-independent mortality rate (expressed as a *k*-value) of *G. m. submorsitans* in the Yankari Game Reserve, Nigeria (averages for the years 1964–74) and the mean saturation deficit of the previous month (averages from within the period 1941–70), $y = 0.051 + 0.011x$, $r = 0.914$, $P < 0.0001$; b, *G. m. submorsitans* (as in Fig. 1a) and the mean monthly NDVI of the Yankari Game Reserve, Nigeria, of the previous month (averages for the years 1985–87), $y = 0.442 - 0.623x$, $r = 0.893$, $P < 0.001$; c, *G. fuscipes fuscipes* near Lugala, Uganda (averages for the years 1971–76) and the mean monthly NDVI of the previous month (averages for the years 1984–88), $y = 0.501 - 0.612x$, $r = 0.656$, $P < 0.05$. NDVI = (channel 2 – channel 1)/(channel 2 + channel 1), here and later (except where stated) read from composited monthly images on a Hammer-Aitoff projection with a resolution of 7.6 km. The number against each point refers to the month, *n* (1 = January, and so on).

shown in Fig. 1, is about 0.3 (equivalent to a maximum rate of population increase of twofold per month). Therefore, areas of Africa characterized by an annual saturation deficit associated with a mortality in excess of this maximum rate of increase should be unsuitable for this species of fly. Within the distribution limits of the flies, abundance must be a function of the difference between the annual fertility and annual mortality of the flies.

To predict the likely occurrence and abundance of tsetse far from the original study site requires accurate meteorological information, but there are only 362 meteorological stations in the area of Africa affected by tsetse⁹ (latitudes 15° N–25° S), or one per 28,000 km². Recently, rainfall in the Sahel has been correlated with reflectance in the visible and infrared wavebands of the National Oceanic and Atmospheric Administration (NOAA) series of meteorological satellites (resolution down to

1 km)¹⁰ and with cloud-top temperatures from Meteosat data (resolution down to 5 km)¹¹. The channel 1 (red) and channel 2 (near infrared) data of the Advanced Very High Resolution Radiometer (AVHRR) of the NOAA satellites are also used to produce normalized difference vegetation indices (NDVIs) related to the intercepted photosynthetically active radiation (IPAR)¹² and vegetation type on continental scales¹³. These are the most widely available satellite data and are important predictors of vegetation biomass¹⁴, seasonal crop production¹⁵ and the likelihood of famines¹⁶. A natural extension of this is to use weighted vegetation indices to detect the ephemeral presence of green vegetation in or near locust outbreak and invasion areas and therefore the potential for locust development¹⁷. High weighted NDVI values are also associated with outbreaks of Rift Valley Fever virus in Kenya¹⁸, owing to seasonal breeding of the mosquito vectors. In addition, the much higher spatial resolution LANDSAT (30 m) or SPOT (10 m) data have been used to characterize habitats of other mosquitoes¹⁹ and ticks²⁰, trematodes²¹, tsetse²² and dunlin²³.

We investigated the utility of satellite data in studies of tsetse by examining correlations between mean annual NDVIs and the mean annual values of temperature, saturation deficit and rainfall for sites throughout the tsetse zone of Africa²⁴. The results (Fig. 2) show that NDVIs are significantly negatively correlated with saturation deficit, positively, but nonlinearly, correlated with rainfall²⁵, but not correlated with temperature. Figure 1b shows the correlation between the mortality rate of *G. morsitans* (as in Fig. 1a) and the NDVI of the previous month in Nigeria. Figure 1c shows a similar correlation for an East African species of the forest and riverine group of tsetse, *G. fuscipes*, from Uganda.

G. palpalis, the widespread vector of human sleeping sickness in West Africa, was studied along a 700-km transect in Côte d'Ivoire and Burkina Faso in the wet and dry seasons of 1983–84. Human trypanosomiasis is a major problem only in the middle of the area sampled by the transect. Populations of *G. palpalis* were sampled at each of eight sites using biconical traps²⁶ and later analysed in standard ways to estimate size, age and nutritional condition²⁷. Figure 3 records the relationship between the mean length of the hatchet-cell vein in the wing of male and

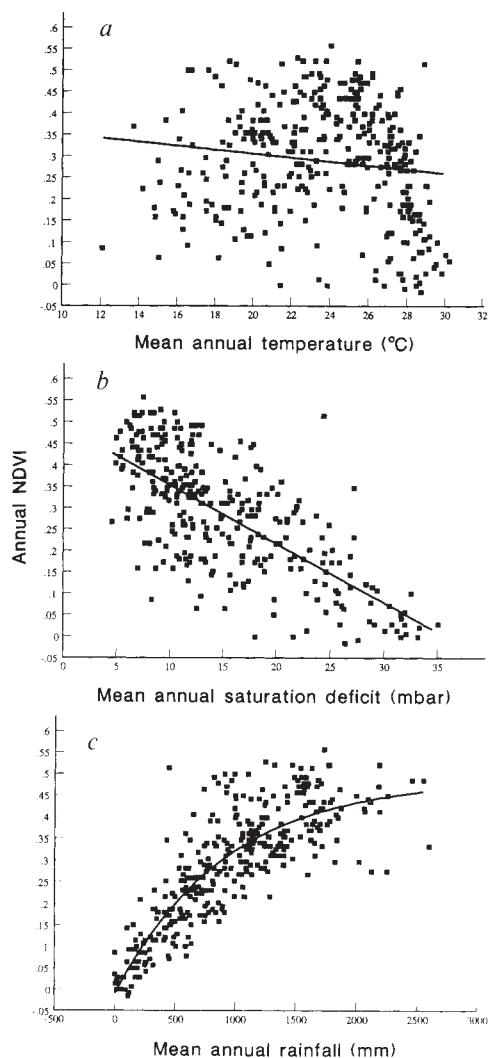


FIG. 2 The relationship between the annual NDVIs (for the early 1980s) and meteorological conditions at sites throughout Africa below latitude 20° N (for the period 1941–70). a, Mean annual temperature, $y = 0.397 - 0.0046x$, $r = 0.13$, not significant; b, mean annual saturation deficit $y = 0.489 - 0.0137x$, $r = 0.700$, $P < 0.001$; c, mean annual rainfall, $y = -0.01 + 0.497(1 - e^{-0.0011x})$, the second two parameters, but not the first, are significantly different from zero at the 5% level. All sites recorded in HMSO 1983 (ref. 9) were included in this analysis except coastal stations (for which image registration was occasionally poor and where human activity associated with ports makes the interpretation of NDVIs problematic). Point data from Mercator projection NDVIs with a resolution of 25 km.

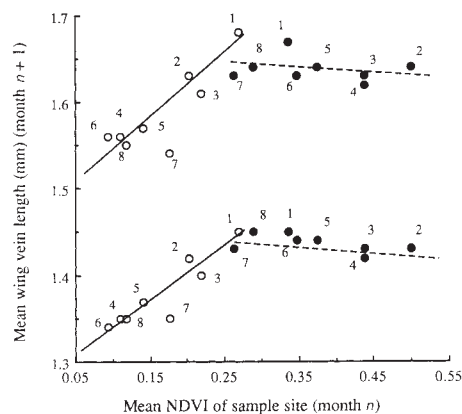


FIG. 3 The relationship between the mean length of the hatchet cell vein (a measure of the size of the fly) in populations of *G. palpalis* at eight sites along a ~700-km north-south transect in West Africa and the NDVI of the previous month at each sample site in the wet (●) and dry (○) seasons of 1983 and 1984 respectively. Upper traces, females (dry season, $y = 1.476 + 0.670x$, $r = 0.852$, $P < 0.01$; wet season, $y = 1.651 - 0.037x$, $r = 0.201$, not significant). Lower traces, males (dry season, $y = 1.279 + 0.598x$, $r = 0.915$, $P < 0.01$; wet season, $y = 1.467 - 0.086x$, $r = 0.584$, not significant). The sample sites, 100 km apart, are sequentially from south to north: 1, Pauli Brousse, near Sassandra, on the coast of Côte d'Ivoire; 2, Antonihio; 3, Degbèzère; 4, Bo'Pri; 5, River N'zi; 6, Komborodougou; 7, Oulokoussou; 8, La Guingette, near Bobo-Dioulasso, Burkina Faso.

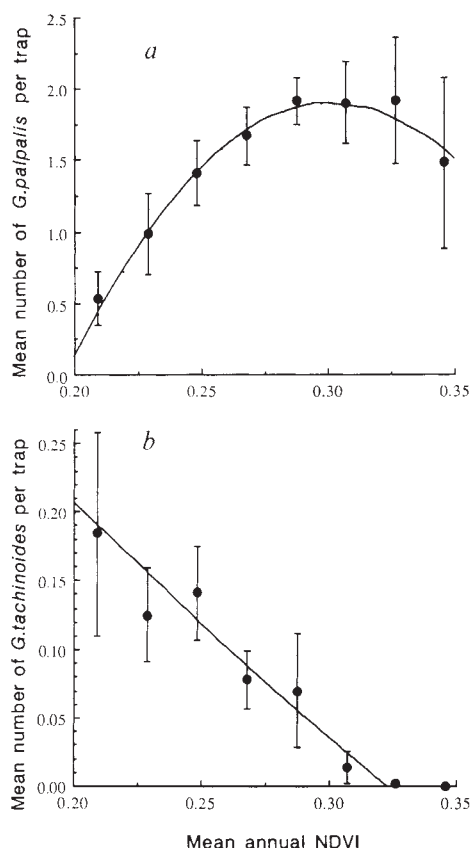


FIG. 4 The relationship between the mean (± 1 s.e.) daily biconical trap catches of *G. palpalis* (a) and *G. tachinoides* (b) and mean annual NDVI within each 0.25° (~ 28 km) square of the FAO/GTZ survey of northern Côte d'Ivoire²⁷. For each of the 208 sites the NDVI is the mean value for 16 equally spaced pixels within each 0.25° square, for the period August 1981–July 1990. The mean fly density for each site was then averaged within successive 0.02 NDVI units, for example NDVI 0.20–0.219, 0.22–0.239. The four survey sites within the Comoe National Park are omitted as fly densities were abnormally high, probably because of high host densities within the park. The fitted lines are second order polynomials: $y = -13.857 + 104.63x - 173.66x^2$, $n = 208$, $r = 0.257$, $P < 0.001$ (a); $y = 0.625 - 2.332x + 1.228x^2$, $n = 208$, $r = 0.239$, $P < 0.001$ (b).

female flies (a measure of fly size) and the NDVIs of the sample site the month before the samples were collected. The size of flies is determined during larval development by the conditions experienced by the pregnant female about one month before the samples were taken. Whereas there is no change in mean vein length across the transect in the wet season, when NDVIs are relatively high, there is a significant relationship between vein length and NDVI in the dry season when the more northerly sites have smaller flies (Fig. 3). For *G. pallidipes* in Kenya, a decrease in vein length of 0.05 mm was associated with an increase in fly mortality rate of 0.5 on the k -value scale²⁸. Along the West African transect there are relatively higher density, less stressed fly populations in the South and lower density, climatically more stressed populations in the North. In the middle of the transect zone, fly populations of intermediate density may bite enough humans (not a preferred host) during the seasonally stressful dry season to create the greatest epidemiological risk.

The most fully documented survey of tsetse abundance was performed in 1979–80 in the northern half of Côte d'Ivoire²⁹ (that is, throughout the savannah region), an area of

140,000 km². Figure 4 shows the relationship between the numbers of *G. palpalis* and *G. tachinoides* per trap per day throughout the study zone and NDVI values. *G. palpalis* numbers peak in the mid to high part of the NDVI range (Fig. 4a), whereas those of *G. tachinoides* (a species adapted to drier conditions) are highest at the low end of the NDVI range of the sample sites (Fig. 4b). The relationship between animal abundance and climate (or NDVI) arises because of the density dependence in the system⁸, which has been demonstrated for tsetse³⁰. This analysis suggests that remote sensing can exploit such a feature of biological systems, locally or regionally, to predict the levels of abundance of tsetse and, by implication, many other pests and disease vectors.

A long-term aim of the present study is to produce maps of high risk of disease transmission for the very large areas of the tropics where vector-borne diseases greatly affect human and animal welfare. That remote sensing provides a means by which we may assess two key parameters of vector-borne diseases—vector mortality rate and abundance³¹—is a significant step towards this goal. □

Received April 26; accepted 21 May 1991.

- Murray, M. & Gray, A. R. *Prev. vet. Med.* **2**, 23–30 (1984).
- De Mynck, A. & Rogers, D. J. (eds) *Ann. Soc. Belg. Med. trop.* **69**, (1989).
- Mulligan, H. W. (ed.) *The African Trypanosomiasis* (Allen & Unwin, London, 1970).
- Rogers, D. J. & Randolph, S. E. *J. Anim. Ecol.* **55**, 1007–1025 (1986).
- Moran, P. A. P. *Biometrics* **6**, 250–258 (1950).
- Buxton, P. A. *London School of Hygiene & Tropical Medicine, Memoir 10* (Lewis, London 1955).
- Rogers, D. J. *J. Anim. Ecol.* **48**, 825–849 (1979).
- Varley, G. C., Gradwell, G. R. & Hassell, M. P. *Insect Population Ecology* (Blackwell Scientific, Oxford, 1973).
- HMSO *Tables of Temperature, Relative Humidity, Precipitation and Sunshine for the World, Part 4* (HMSO, London, 1983).
- Barrett, E. C. & D'Souza, G. *Proc. 20th Int. Symp. Rem. Sens. Env., Nairobi, Kenya* **3**, 955–964 (1986).
- Milford, J. R. & Dugdale, G. *Rainfall mapping over West Africa in 1986* (Dept Meteorology, Univ. Reading, 1986).
- Tucker, C. J. & Sellers, P. J. *Int. J. Rem. Sens.* **7**, 1395–1416 (1986).
- Tucker, C. J., Townshend, J. R. G. & Goff, T. E. *Science* **227**, 369–375 (1985).
- Huetten, A. R. & Jackson, R. D. *Rem. Sens. Env.* **23**, 213–232 (1987).
- Bartholomé, E. *Int. J. Rem. Sens.* **9**, 1539–1552 (1988).
- Hutchinson, C. F. *Int. J. Rem. Sens.* (in the press).

- Hielkema, J. U., Roffey, J. & Tucker, C. J. *Int. J. Rem. Sens.* **7**, 1609–1622 (1986).
- Linthicum, K. J., Bailey, C. L., Davies, F. G. & Tucker, C. J. *Science* **235**, 1656–1659 (1987).
- Washino, R. K., Pitcairn, M. J. & Rejmankova, E. *Mosquito Control Research 1987 A. Rep.* (Univ. California, Division of Agriculture & Natural Resources, 1987).
- Hugh-Jones, M. E. *et al. Acta Vet. Scand.* **84** (Suppl.), 256–261 (1988).
- Cross, E. R. & Bailey, R. C. *Milit. Med.* **149**, 28–30 (1984).
- Bourn, D. thesis, Oxford Univ., 1983.
- Avery, M. I. & Haines-Young, R. H. *Nature* **344**, 860–862 (1990).
- Ford, J. & Katondo, K. M. *The Distribution of Tsetse Flies in Africa* (OAU/Cook, Hammond & Kell, London, 1977).
- Nicholson, S. E., Davenport, M. L. & Malo, A. R. *Climatic Change* **17**, 209–241 (1990).
- Challier, A. & Laveissière, C. *Cah. ORSTOM Sér. Entomol. méd. Parasitol.* **11**, 251–262 (1973).
- Randolph, S. E. & Rogers, D. J. *Bull. ent. Res.* **68**, 655–671 (1978).
- Dransfield, R. D., Brightwell, R., Kiilu, J., Chaudhury, M. F. & Adabie, D. A. *Med. vet. Ent.* **3**, 83–95 (1989).
- FAO/GTZ *Cartographie de la Répartition des Glossines* (IEMVT, Maisons Alfort, France, 1982).
- Rogers, D. J. & Randolph, S. E. *Insect Sci. Applic.* **5**, 397–402 (1984).
- Rogers, D. J. *Parasitology* **97**, 193–212 (1988).

ACKNOWLEDGEMENTS. We are grateful to J. Townshend, J. Hielkema and G. Farmer for the supply of images used in this analysis and B. Williams for comments on the manuscript.

Cloning of a putative high-affinity kainate receptor expressed predominantly in hippocampal CA3 cells

Pia Werner, Mark Voigt, Kari Keinänen*, William Wisden & Peter H. Seeburg†

Laboratory of Molecular Neuroendocrinology, Centre for Molecular Biology, University of Heidelberg, Im Neuenheimer Feld 282, 6900 Heidelberg, Germany

* Present address: VTT, Biotechnical Laboratory, PO Box 202 (Tietotie 2), SF-02150 Espoo, Finland

† To whom correspondence should be addressed

KAINIC acid is a potent neurotoxin for certain neurons¹. Its neurotoxicity is thought to be mediated by an excitatory amino-acid-gated ion channel (ionotropic receptor) possessing nanomolar affinity for kainate². Here we describe a new member of the rat excitatory amino-acid receptor gene family, KA-1, that has a 30% sequence similarity with the previously characterized α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor subunits GluR-A to -D³⁻⁵. The pharmacological profile of expressed recombinant KA-1 determined in binding experiments with [³H]kainate is different from that of the cloned AMPA receptors and similar to the mammalian high-affinity kainate receptor (kainate > quisqualate > glutamate > AMPA) with a dissociation constant of about 5 nM for kainate^{2,6}. The selectively high expression of KA-1 messenger RNA in the CA3 region of the hippocampus closely corresponds to autoradiographically located high-affinity kainate binding sites⁷⁻⁹. This correlation, as well as the particular *in vivo* pattern of neurodegeneration observed on kainate-induced neurotoxicity^{1,2}, suggests that KA-1 participates in receptors mediating the kainate sensitivity of neurons in the central nervous system.

In a search for further members of the ionotropic glutamate

receptor family, a novel full-length complementary DNA sequence was isolated that encodes a predicted 956-residue polypeptide of relative molecular mass about 105,000 ($M_r \sim 105K$). This molecule (KA-1) has sequence similarities (Fig. 1) to the AMPA receptor subunits GluR-A to -D (refs 3-5) as well as to kainate-binding proteins (KBPs) of chick¹⁰ and frog¹¹ in its C-terminal half (KA-1 residues 400-800, 53% identity to GluR-A and 37% to chick KBP). The region of sequence similarity includes the four hydrophobic regions important for transmembrane topology and channel formation in GluR-A to -D^{3,5,12}. The N-terminal portion of the KA-1 polypeptide, of predicted extracellular location and containing consensus N-linked glycosylation sites, shows very low similarity to the corresponding regions in the AMPA receptor subunits, and none to the much shorter extracellular regions of the two KBPs.

KA-1 transiently expressed in cultured cells^{3,13} was analysed for its ability to bind ligands of the excitatory amino acid (EAA) receptor family. No high-affinity [³H]AMPA binding was detected (not shown), thus clearly distinguishing this molecule from subunits of the AMPA receptor family. [³H]kainate bound with a dissociation constant (K_d) of 4.7 ± 0.7 nM ($n = 3$) in a reversible and saturable manner (Fig. 2), in close agreement with K_d values for the high-affinity binding of this ligand in brain membranes¹⁴ and by *in vitro* receptor autoradiography⁹. Although studies using brain tissue have identified two high-affinity sites for [³H]kainate (K_d of 5 nM and 50 nM)^{9,14}, we observed only one site on cells expressing KA-1.

In competition experiments using EAA analogues, KA-1 had the pharmacology typical of a high-affinity kainate receptor, with a rank order of potency of kainate > quisqualate > glutamate > AMPA^{6,14,15}. The inhibitory constants (K_i) for quisqualate, L-glutamate and AMPA were about 18,200 and >5,000 nM, respectively. Hence, KA-1 has marked differences in its pharmacology to either the chick KBP (K_d for kainate of 0.6 μ M¹⁶) or the amphibian KBP, at which kainate and AMPA have equal affinities¹⁷. Domoate had a K_i of 40 nM at KA-1, markedly lower than the K_i of 1.5 nM for the kainate site with a K_d of roughly 5 nM in brain¹⁴. This may indicate a molecular

FIG. 1 Polypeptide sequence of KA-1 and comparison with AMPA receptor GluR-A flop³ and chick kainate binding protein¹⁰. Only substitutions to KA-1 are shown. Dotted positions represent artificial gaps introduced for better alignment. All protein sequences are numbered starting with their predicted mature N termini²⁸ (single-letter amino-acid code). The C termini are indicated by stars. Signal sequences are underlined; putative transmembrane regions TM1 to TM4 are boxed; ●, all possible N-linked glycosylation sites in the predicted extracellular domain of KA-1; +, consensus phosphorylation sites for Ca²⁺-calmodulin-dependent protein kinase type II (S/T-(x-x)-D/E, where x is an unknown amino acid) in the predicted intracellular domain of KA-1.

METHODS. Partial KA-1 sequence was obtained from rat brain cDNA using the polymerase chain reaction (PCR). Oligonucleotide primers and reaction conditions for PCR were as described³. The amplified DNA (~520 base pairs) was gel-excised and cloned into M13 vectors for DNA sequence analysis, using the *Eco*RI and *Kpn*I sites engineered into the

KBP (Chick)	DKGLHFIFCVTVAV LLRE SQ.
GluR-A (Flop)	YIF FECTGLGAVVG NF NNIQ GGLFNPQOS...Q HAAFRF LSQTEP KL PQIDI...VNISDTF MYRF SQFS Y IF FYER
KA-1	MPRVSAPIVLLPAILMLVACSPHSLRIAAILDDPECSRGERLSITLAKNRINRAPELKGAKVEVDIFELLRLDSEYETAETMCQILPKGVAVLG...P
74	RTVNMILT FOGALHVCFT S P D TSNQFVL R.....PELQE LIS IDHYKQ FVY YDADRG SV QRW DTAEEKNMQVTA NI TT
78	SSSPASSIISNICGEKVEPHKVAPEEFVRFQLOQTTLNLHPSNTDISVAVAGILNFFNCTACILCAKAECLLNLEKLLRQLFSK. DTLSVRMLDD
163	EEGYRM FQDLEKK ERLVVVDCESERLNA GQIVK EKNGIG H LA G MDIDLNKFESGR VT QLV YDTIPARIM QWRT DSRDHT
177	TRDPTPLL.KEIRDDKTATIIHANASMSHTILLKAEELGWSAYTYTYIFNLEFSLQMDSLVDDRNVNIGFSIFNQSHAFFQEFQSQSLNQSQWQENCDH
263	DNKR KYT TY G KVMAE F S R QRDISRRGNA DCLANP VP GQ IDIQRA QQ RF NVQ E R T T HVIEKMD I K
276	VPFTGPALSSALLFDVAVVAVTAVQELNRSQEIQVGLPSCG...SAQIQHGTSLMNYLRVLEGLTGHIEFNSKGRSNYALKILOFTNRGFRQIG
1	...TGAMRND AMIKENDLRGPEE LPS T D V VRSAE.....L Y I L A SM HS VKV K AISP.S N I
363	Y NEDDKFVPAATDAQ GGDNS VQ R YI D V K ANQF Y ELAA I KHVGS RLEI S K ARDPDTKA N
371	QWVHAEL. SMDSRLYASNISDSIFN. TTLVVTITLENPIYMLKGNHQDMGNDREYEGFVDMLEKELALIRFNKYIRLVGDGVYGP/PEANG. TWGTMVG
88	ILRQE I P V SA E VS TT LQT G L KETISQEMSF HF A KET TGL F VLTVC S C NE KNEE.....
462	VYGR V P LV E S MIKPK.QKSK V F LAYEL MCIVF IG V S FS H EEFEEGRDQT
468	ELIARKADLAVAGLTIITAREKVIDFSKPFMTLGISILYRVH.MGRRPGYFSSLDFFSPQVLMFLIAYLAVSCVFLVLRILTPEYEWSP...HPCAGGRC
182	...HTTFL GA ALTL V PR K F V VIAAI L TALLAA I FT L SSGSEQLS QTTE VK RKL F LD S TFY K KN
561	SDQS EFGIF SLA CD S S G I G F E VS AE K E A LEA TKE RR KI
565	NLLVNYQYSLGNSLFPVVGFMQGSITAPRALSTRCVSGVWMAFTLIISSYTNIAAFLTVORMVEIETSDVLDLADQTAIEYGTIHGSSMTFFQNSRY
278	PIHRMVE DKRRDL L TYQ AVQ ME .. IG ISQDLAA H. IRAPEVIGAR F ATAQA PWTKKLSV V K R TGD DY RN
661	AVEFK T K AE RT MI RK KGK Y IE KP DTMKV N S AT K AL NPVN V K N QGL DK N
665	QTYQRMWNYMSKQPSVFVKSTEEGIARVLNS..NYAFLESTMEYRYQRN.CNLTQIGGLDITKGYIGMPVGSVFRDEFDLAILQLQENNRLEIKR
375	...ESS LHKSEGWSP QPQAL L LT AI ALGVIA V LSNKS AAGHIKKS CSIF EMCTR RIKENTROTQETS RANA*
761	YDKG C SGGGSKDKTSA SLVS VA V YI G GL MLV LI CYKSSESK MKGFC I QSIN
762	KWV...EGGKCPKEEDHRAKGLGMEIGGIFVVLICGLIVAIFMAIEFLIWTLRHSEASEVSCQEMMTELRSLIILQDNIHPRRRSSGGLPQPVPVLE
835	AI TSTLPRN GAGAS G GSGENGRVVS DFFKSMQSPICMSSH S M LGATGL*
858	ERRPRGTATLSNGKLCGAGEPDLAQLRAQEAALVARGWHAHPRVGPVLPALPGAGSTVSAQRGEPGVGDHQQQRA*

PCR primers. A novel sequence related to that of AMPA receptors GluR-A to -D was identified and its full-length cDNA was isolated from rat brain hippocampal cDNA libraries constructed in the λ gt10 vector, using the ³²P-labelled cloned PCR fragment as a probe.

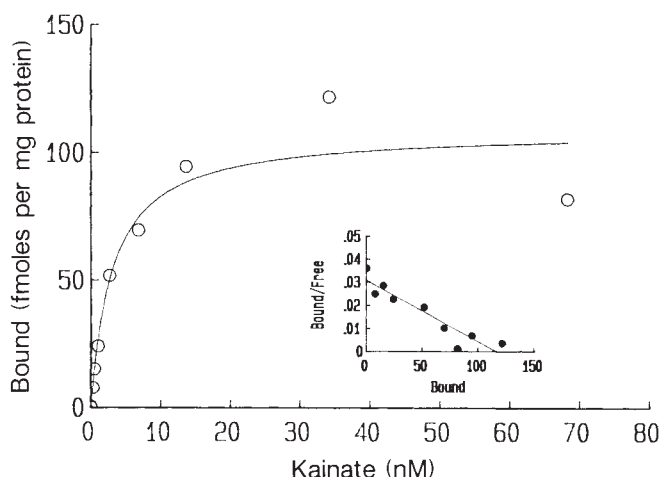


FIG. 2 Saturation analysis of [3 H]kainate binding to KA-1. Binding was performed using ligand concentrations of 0.2–100 nM. Results were analysed using a nonlinear regression program (Graphpad), and were best-fitted to a single-site model. The experiments shown yielded a K_d of 3.9 nM and a maximum binding (B_{max}) of 116 fmol per mg protein.

METHODS. A 3-kbp *Clal*–*XhoI* DNA fragment containing the entire coding region of KA-1 was subcloned into a eukaryotic expression vector^{3,13}. Human embryonic kidney 293 cells (ATCC CRL 1573) were transfected with this construct^{3,29}, and the cells collected 48 h post-transfection. Cells were washed twice with $1 \times$ phosphate-buffered saline, removed from the plate, centrifuged at 15,000g, and resuspended into 50 mM Tris-citrate, pH 7.0. After three more rounds of homogenization and centrifugation, the membranes were brought up in 50 mM Tris-citrate, pH 7.0 to their final volume at a concentration of about 30 μ g protein per assay tube. Binding assays were carried out in a total volume of 0.5 ml for 60 min at 0°C using [3 H]kainate (58 Ci mmol⁻¹, NEN), with nonspecific binding defined in the presence of 1 mM L-glutamate. Competition studies were performed in the presence of 5 nM [3 H]kainate. Mock-transfected cells exhibited no specific binding at all concentrations of [3 H]kainate tested.

heterogeneity of the mammalian high-affinity kainate site and/or a missing subunit for its recombinant reconstitution.

Brain slice studies have demonstrated that kainate at nM concentrations increases neuronal excitability^{18–20}. We did not obtain kainate- or glutamate-evoked currents in cells expressing KA-1, which may reflect very small unitary channels, low open probability or fast desensitization kinetics of the homooligomeric KA-1 assemblies. Indeed, absence of homooligomeric-channel activity has been reported for subunits of other ligand-gated ion channels (for example, ref. 21). In the EAA receptor family, no channel activity was observed for the KBPs^{10,11} or for a recently cloned member of unknown function²². No consistent change in current properties was seen on co-expression of KA-1 with AMPA receptor subunits, suggesting that these molecules may not be natural channel partners. But the presence of the core structure conserved in the AMPA receptor subunits, including a region with functional channel

signature in the putative second transmembrane region¹², would predict that KA-1 participates in the formation of cationic channels.

We examined the distribution of the KA-1 messenger RNA through horizontal, sagittal and coronal planes of the adult rat brain by *in situ* hybridization²³. The KA-1 gene is expressed strongly in the CA3 region of the hippocampus. By comparison, low mRNA levels are seen in hippocampal dentate granule cells; in layers II, V and VI of cortex; and in cerebellar Purkinje cells (Fig. 3d). No mRNA could be detected in the striatum, reticular thalamus, hypothalamus or amygdaloid complex. Cellular resolution of hippocampal cells indicates that the KA-1 mRNA is confined to CA3 pyramidal cells and is largely absent from the CA1 cells (Fig. 3e,f).

This distribution of KA-1 mRNA is similar to that obtained from high-affinity [3 H]kainate autoradiography in rat brain^{7–9}, where abundant binding is seen in layers I, V and VI of cortex,

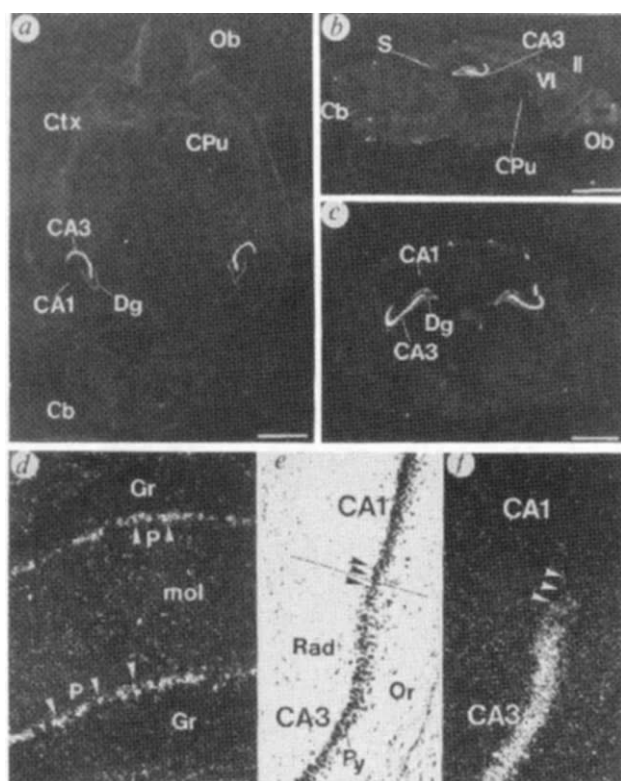


FIG. 3 Expression of KA-1 mRNA in the adult rat brain. *a–d*, X-ray film autoradiographs demonstrating prominence of KA-1 mRNA in the hippocampal CA3 area. *a*, Horizontal section; *b*, sagittal section; *c*, coronal section; *d*, dark-field photomicrograph of emulsion-dipped section, showing accumulation of silver grains over cerebellar Purkinje cells (arrowheads, P); *e*, *f*, bright-field, dark-field optics of cellular expression of KA-1 mRNA in the hippocampus. There is a large amount of KA-1 mRNA in CA3 pyramidal cells, but a virtual absence in those of CA1. Arrows and dotted line in *e* mark transition from large, loosely packed CA3 pyramidal cells to small, densely packed CA1 cells. Cb, cerebellum; CPu, caudate putamen; Ctx, cortex; Dg, dentate granule cells; Gr, cerebellar granule cells; mol, molecular layer; Rad, stratum radiatum; S, subiculum; II, VI, layers of neocortex. Scale bars: *a*, 3 mm; *b*, 5 mm; *c*, 2.5 mm; *d*, 150 μ m; *e*, 280 μ m.

METHODS. *In situ* hybridization with 35 S-labelled antisense oligonucleotides (45 and 42 mers) was performed exactly as described^{23,30}. Specificity was determined by incubating radiolabelled probes in the presence of a 50-fold excess of unlabelled oligonucleotide. This resulted in blank autoradiographs. As a further check on probe specificity, an identical hybridization pattern was made using three antisense oligonucleotides constructed to different parts of the KA-1 mRNA. These oligonucleotides were 5'-CTTGAGT-TGAACCGTAGGATCTCAGCGAACTCCTTGAGCATGTC-3', complementary to codons for KA-1 amino-acid residues 430–444; 5'-TAGCCCGGTCTGCG-TCCATATGAACCTCTGTAAGAATACTA-3', complementary to codons for residues 503–516; 5'-GTTGACCAGGAGATTACACCGCCCTGTGCACAAG-GATGTGGACT-3', complementary to codons for residues 555–569. All images shown were obtained with the first oligonucleotide.

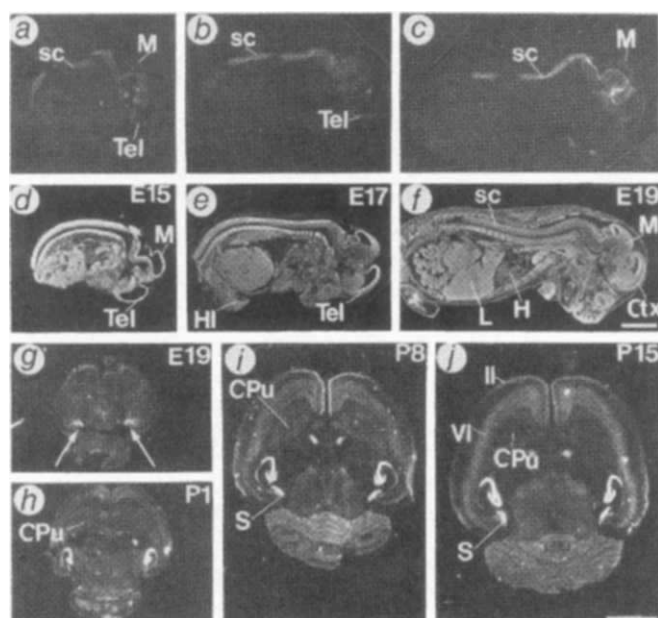


FIG. 4 Developmental expression of KA-1 mRNA. *a-f*, KA-1 mRNA in E14 (*a, d*), E17 (*b, e*) and E19 (*c, f*) in whole embryos; *a-c* are the X-ray film images; *d-f* correspond to Nissl stains of similar sections. Abbreviations: Ctx, cortex; CPu, caudate putamen; II, VI, layers of neocortex; H, heart; HI, hindlimb; L, liver; M, mesencephalon; sc, spinal cord; Tel, telencephalon; scale bar in *f*, 3.7 mm. *g-j*, Horizontal sections through the developing rat brain; *g*, E19; *h*, P1; *i*, P8; *j*, P15; arrows in *g* indicate KA-1 mRNA in the hippocampal formation; S, subicular complex; scale bar in *j*, 4 mm. KA-1 mRNA is present throughout the embryonic spinal cord and in restricted areas of presumptive hindbrain and forebrain. KA-1 mRNA is not detected in the liver/heart/gut compartments or in bone.

and in CA3 and dentate gyrus of hippocampus. By contrast, CA1 cells show few such sites. This expression pattern of KA-1 correlates well with the selective vulnerability of CA3 neurons to kainate^{1,2}. But regions of brain that appear to lack KA-1 mRNA, including the reticular thalamus, striatum and cerebellar granule layer^{8,9}, contain high-affinity kainate binding sites and therefore may express receptors containing KA-1-related subunits.

Mammalian high-affinity kainate receptors may reside on presynaptic sites, as suggested by receptor autoradiography and neurotoxicity studies^{2,27}. For example, an intact and mature mossy-fibre projection to the pyramidal CA3 neurons is needed for high levels of kainate binding²⁷ and for kainate-mediated neurotoxicity²⁴ in this region. A presynaptic localization of kainate sites is supported by the localization of KA-1 mRNA in dentate granule cells, which are the source of the mossy fibres. But our finding regarding the high expression of KA-1 transcripts by CA3 neurons suggests that these cells contribute prominently to the kainate receptor population in the CA3 region.

EAA transmitter activity plays an important part in the construction of circuitry in the developing brain²⁴. When examining the expression of KA-1 mRNA in the embryonic and early postnatal brain, transcripts could be detected as early as E15 (embryonic day 15) in brain and spinal cord (Fig. 4*a, d*). By comparison, very low levels were detected in mesencephalon and in the telencephalic areas destined to become cortex²⁵. This

pattern remains essentially invariant until E19 (Fig. 4*c, f*). By this time, a heavy accumulation of the mRNA is evident in the hippocampal formation (Fig. 4*g*, arrows), even though this structure has not yet completely formed and cells are still migrating into position²⁵. By postnatal day 1 (P1), the CA3 and dentate granule cell expression seems well established (Fig. 4*h*). There are three apparent developmental traits of KA-1 mRNA expression in postnatal brain: (1) there is a prominent expression of KA-1 mRNA in the subicular cortex at P1, P8 and P15, but this is markedly reduced in the adult (Figs 4*h-j* and 3*b*); (2) levels of KA-1 mRNA in the striatum decrease from P1 when KA-1 mRNA is equibundant in cortex and striatum (Fig. 4*h*) through P15, when striatal KA-1 mRNA becomes undetectable (Fig. 4*j*); (3) the cortical expression pattern is also more pronounced in animals of P8 and P15 stages (Fig. 4*i, j*) compared with the adult (Fig. 3*a-c*). These observations agree with developmental autoradiographic studies²⁶ reporting high levels of [³H]kainate binding in the inner lamina of neocortex in the first three postnatal weeks, with lower levels thereafter. By contrast, the qualitative pattern of KA-1 mRNA expression remains unchanged throughout postnatal development, also consistent with that study²⁶.

KA-1 fulfils the criteria for a subunit of a mammalian high-affinity kainate receptor by its structural signature, pharmacology and pattern of expression in brain. The availability of this subunit for a putative high-affinity kainate receptor should aid in the characterization of the cognate *in vivo* site. □

Received 2 April; accepted 14 May 1991.

- Nadler, V. J., Perry, B. W. & Cotman, C. W. *Nature* **271**, 676-677 (1978).
- Coyle, J. T. *J. Neurochem.* **41**, 1-11 (1983).
- Keinänen, K. *et al. Science* **249**, 556-560 (1990).
- Boulter, J. *et al. Science* **249**, 1033-1037 (1990).
- Nakanishi, N., Shneider, N. A. & Axel, R. *Neuron* **5**, 569-581 (1990).
- Young, A. B. & Fagg, G. E. *Trends pharmac. Sci.* **11**, 126-133 (1990).
- Foster, A. C., Mena, E. E., Monaghan, D. T. & Cotman, C. W. *Nature* **289**, 73-75 (1981).
- Monaghan, D. T. & Cotman, C. W. *Brain Res.* **252**, 91-100 (1982).
- Unnerstall, J. R. & Wamsley, J. K. *Eur. J. Pharmacol.* **86**, 361-371 (1983).
- Gregor, P., Mano, I., Maoz, I., McKeown, M. & Teichberg, V. *Nature* **342**, 689-692 (1989).
- Wada, K. *et al. Nature* **342**, 684-689 (1989).
- Verdoorn, T. A., Burnashev, N., Monyer, H., Seeburg, P. H. & Sakmann, B. *Science* (in the press).
- Gorman, C. M., Gies, D. R. & McCray, G. *DNA Prot. Engin. Tech.* **2**, 3-9 (1990).
- Hampson, D. R., Huie, D. & Wenthold, R. J. *J. Neurochem.* **49**, 1209-1215 (1987).
- Honoré, T. *et al. Science* **241**, 701-703 (1988).
- Henley, J. M. & Barnard, E. A. *J. Neurochem.* **56**, 702-705 (1991).
- Henley, J. M., Ambrosini, A., Krosgaard-Larsen, P. & Barnard, E. A. *New Biol.* **1**, 153-158 (1989).
- Robinson, J. H. & Deadwyler, S. A. *Brain Res.* **221**, 117-127 (1981).
- Westbrook, G. L. & Lothman, E. W. *Brain Res.* **273**, 97-109 (1983).

- Ben-Ari, Y. & Gho, M. *J. Physiol., Lond.* **404**, 365-384 (1988).
- Deneris, E. S., Connolly, J., Rogers, S. W. & Duvoisin, R. *Trends pharmac. Sci.* **12**, 34-40 (1991).
- Bettler, B. *et al. Neuron* **5**, 583-595 (1990).
- Wisden, W., Morris, B. J. & Hunt, S. P. *Molecular Neurobiology: A Practical Approach* Vol. 2 (eds Chad, J. & Wheal, H.) 205-225 (Oxford University Press/IRL, Oxford, 1991).
- McDonald, J. W. & Johnston, M. V. *Brain Res. Rev.* **15**, 41-70 (1990).
- Paxinos, G., Törk, I., Tecott, L. H. & Valentino, K. L. *Atlas of the Developing Rat Brain* (Academic, San Diego, 1990).
- Miller, L. P., Johnson, A. E., Gelhard, R. E. & Insel, T. R. *Neuroscience* **35**, 45-51 (1990).
- Represa, A., Tremblay, E. & Ben-Ari, Y. *Neuroscience* **20**, 739-748 (1987).
- von Heijne, G. *Nucleic Acids Res.* **14**, 4683-4690 (1986).
- Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745-2752 (1987).
- Monyer, H., Seeburg, P. H. & Wisden, W. *Neuron* **6**, 799-810 (1991).

ACKNOWLEDGEMENTS. We thank H. Monyer and M.-L. Malosio for the embryonic sections, A. Herb, S. Grünwald and U. Keller for help with DNA sequencing, cell culture and *in situ* hybridization, respectively, and J. Rami for secretarial assistance. M.V. is supported by the Alexander von Humboldt Foundation; K.K. held and W.W. holds an EMBO long-term fellowship. This work was supported by the BMT and the Fonds der Chemischen Industrie (P.H.S.). The nucleotide sequence of KA-1 will appear in the EMBL, Gen Bank and DDBJ nucleotide sequence databases under the accession number X59996.

Cloning of a cDNA for a glutamate receptor subunit activated by kainate but not AMPA

Jan Egebjerg, Bernhard Bettler,
Irm Hermans-Borgmeyer & Stephen Heinemann

Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies,
La Jolla, California 92037, USA

FAST excitatory transmission in the vertebrate central nervous system is mediated mainly by L-glutamate. On the basis of pharmacological, physiological and agonist binding properties, the ionotropic glutamate receptors are classified into NMDA (N-methyl-D-aspartate), AMPA (α -amino-3-hydroxy-5-methylisoxazole-4-propionate) and kainate subtypes¹. Sequence homology between complementary DNA clones encoding non-NMDA glutamate receptor subunits reveals at least two subunit classes: the GluR1 to GluR4 class^{2–6} and the GluR5 class⁷. Here we report the cloning and expression of a functional rat glutamate receptor subunit cDNA, GluR6, which has a very different pharmacology from that of the GluR1–GluR4 class. Receptors generated from the GluR1–GluR4 class have a higher apparent affinity for AMPA than for kainate^{3–6}. When expressed in *Xenopus* oocytes the homomeric GluR6 receptor is activated by kainate, quisqualate and L-glutamate but not by AMPA, and the apparent affinity for kainate is higher than for receptors from the GluR1–GluR4 class. Desensitization of the receptor was observed with continuous application of agonist. The homomeric GluR6 glutamate receptor exhibits an outwardly rectifying current–voltage relationship. *In situ* hybridizations reveal a pattern of GluR6 gene expression reminiscent of the binding pattern obtained with [³H]kainate.

The coding region of a glutamate receptor GluR5 cDNA clone was used as a probe to screen a rat cerebellum cDNA library under low-stringency hybridization conditions⁷. We isolated a 4,559-base-pair (bp) cDNA encoding a protein of 884 amino acids. The protein encoded by the Glu6 cDNA exhibits 80% amino-acid sequence identity with GluR5, but less than 40% identity with the GluR1–GluR4 or with the kainate binding proteins cloned from frog⁹ and chicken¹⁰ (Fig. 1). The similarity between the hydropathy profile of the GluR6 subunit and those of the GluR1–GluR5 subunits suggests that these proteins have a similar membrane-spanning topology.

The physiological and pharmacological properties of the homomeric GluR6 ion channel were studied in *Xenopus* oocytes injected with RNA transcribed *in vitro*. In oocytes held at -100 mV, application of kainate and glutamate evoked inward currents that desensitized in the continuous presence of agonist (Fig. 2a). Full recovery from desensitization caused by application of 100 μ M kainate for 30 s required roughly 15 min. Quisqualate activated only small inward currents but attenuated a subsequent kainate-evoked current (Fig. 2a). AMPA (100 μ M) evoked no detectable current (Fig. 2a; $n=6$) and did not antagonize a kainate-evoked current when AMPA and kainate were applied together (Fig. 2b; $n=6$). The AMPA solution used in these experiments did evoke responses in oocytes injected with either hippocampal messenger RNA (data not shown) or GluR1 RNA (Fig. 2c).

Exposure of the injected oocytes to 10 μ M concanavalin A (conA) for 5 min efficiently decreased desensitization¹¹ and allowed agonist-activated currents mediated by the GluR6 receptor to be studied. ConA treatment increased the current elicited by kainate and glutamate to 75–150 times the peak current for equimolar concentrations before the treatment. After conA treatment, the maximal current induced by glutamate

TABLE 1 EC₅₀ and maximal current relative to kainate-evoked current for GluR6 receptor after conA treatment

Agonist	EC ₅₀ (μ M)	Relative maximal current($\pm$ s.e.m.)
Kainate	1.0 (0.8–1.3)	1.00
Quisqualate	11 (10–13)	0.38 $\pm$ 0.03
Glutamate	31 (29–34)	0.56 $\pm$ 0.03

The mean EC₅₀ are based on measurements of 3–6 oocytes. The numbers in parentheses indicate 95% confidence intervals. Relative maximum current = maximum agonist – evoked current / maximum kainate-evoked current.

relative to kainate was 0.56 ± 0.03 and for quisqualate 0.38 ± 0.03 . ConA-treated oocytes injected with GluR6 RNA which responded to kainate did not respond to 100 μ M aspartate ($n=3$), 100 μ M NMDA with 3 μ M glycine ($n=5$), or to 10 – $1,000$ μ M AMPA (Fig. 2d; $n=8$). Further, coapplication of AMPA (100 μ M) had no effect on the kainate-evoked (1 μ M) responses on conA-treated oocytes ($n=6$). This suggests that AMPA acts as an agonist on only a subset of the kainate/quisqualate-sensitive ionotropic receptors.

The dose response curves for activation of the GluR6 receptor were obtained after conA treatment (Fig. 3a). The concentration needed to evoke 50% of the maximal response (EC₅₀) for kainate (1 μ M) is about 35-fold lower than the EC₅₀ for the homomeric GluR1 (EC₅₀ = 35 μ M) receptor^{2,12}. The EC₅₀ for the GluR6 receptor is 75-fold higher for quisqualate and 10-fold higher for glutamate compared with the same agonist at the GluR1 receptor. Thus the order of agonist potency for the homomeric GluR6 receptor is kainate > quisqualate > L-glutamate (Table 1). This is similar to the order of binding affinities measured for quisqualate and glutamate as competitive displacers of kainate on kainate binding sites in isolated brain membranes¹³. This property is clearly distinct from the GluR1 and GluR3 receptors, where the relative apparent affinities are quisqualate > AMPA > glutamate > kainate^{5,6,13}. On the basis on agonist potencies (EC₅₀) therefore, GluR6 can be considered a kainate receptor of the glutamate receptor family.

CNQX acts as a competitive antagonist of non-NMDA receptors in rat brain neurons¹⁴. CNQX blocked both quisqualate (data not shown) and kainate-evoked responses (Fig. 3e) in oocytes injected with GluR6 RNA. The inhibitory effect of 10 μ M CNQX was eliminated at high kainate concentrations, consistent with its competitive mode of action. CNQX at the same concentration resulted in a 3.5-fold shift of the kainate dose–response curve parallel to the curve obtained in the absence of CNQX (Fig. 3a). Considering the competitive action of CNQX at 10 μ M, the inhibitory concentration (K_i) for CNQX was calculated to be 4 μ M. Thus, CNQX is a less potent blocker of kainate responses at GluR6 receptors than at GluR1 receptors (K_i of 0.519 μ M)¹² and kainate receptors derived from forebrain mRNA (K_i of 0.295 μ M)¹⁵ expressed in oocytes.

We examined the current–voltage relationship (I – V) for kainate and glutamate evoked responses in the presence of conA and for kainate in the absence of conA. We found no qualitative differences between conA-treated and untreated oocytes (Fig. 3b). The I – V relationship exhibited a reversal potential of -10 ± 3 mV and an outward rectification. To analyse whether the outward rectification was an intrinsic property of the channel or perhaps due to an activation of endogenous chloride channels activated by a Ca^{2+} flux¹⁶ through the GluR6 ion channel, we recorded the kainate-evoked I – V relationship in a buffer where 95% of the Cl^- ions were substituted by an equimolar amount of methanesulphonate which is known to shift the chloride reversal potential in a positive direction¹⁷. No significant change was observed in the reversal potential (Fig. 3b), implying that if there is indeed a Ca^{2+} flux in Ringer's solution, it is not sufficient to activate a Cl^- current. The substitution of Cl^- with methanesulphonate reduced the current eightfold; this may have

GluR1: MPYIFAFFCTGFLGAVVG ANFNINIQGGIFPNQSSQ RHAAF
 GluR5: MERSTVLIQGLWTRDTSWLLYLFLCYLP QTSFQVLRIGGIFITVNERPVNVEELAF
 GluR6: MKTIISPVLSNLVFSRSIKVLLCLLWIGYSQG TTH VLNFGGIFITVNESQMGAEELAF
 SIGNAL PEPTIDE

25 RFALSOLTE PPKLLPQIDIVNISDTFEMTYRISQSGVYALFGFYERRTVNM
 29 KFAVTSINNRNRTILENTTLTYDIQRINLFDSEASRRACDQIALGVANFGPSSSSVSA
 27 RFVNTINNRNRTILENTTLTYDTOKINLYDSFEASKACDQSLGVAAIFGPSSSSANA

79 LITSECGALFVCFITPSPFVDTSNQFVLQRLPELOE ALISIDHYKQTFVYIYD
 89 VQSICNALEVPHIQTRWKHPVSVDSDIFTYINLYPDYAAISRAVLDLVLVYNNKTVTVVYE
 87 VQSICNALGVPHIQTRWKHVSVDNKSDFVSLYPDFSSISRATLDLVOFFKWKTVTVVYD

133 ADKGLSVLQRLVDTAAEKNWQVTAVNLTITTEGYRMLFODLEKKERLVVDCESERLN
 149 DSTGLIRLOE LIKAPSRYNIKIKIROLPANKDAKPLLEKMKSKKEFYVIFDCSHETAA
 147 DSTGLIRLOE LIKAPSRYNIRIKIROLPADTKDAKPLLEKMKRGKEFVIFDCSHEMAA

193 AITGQIVKLEKNIGYHYILANTGFMDDINKFKESGRNVTFGDTVNYTITIPARTMOON
 208 EILKQILFMGMTEYYHYFFTTDLDFALDLELYRYSGVNMTGFRLLNIDNPHVSSIIIEKW
 206 GILKQALAMGMYEYYHYIFTTDLDFALDVEFYRYSGVNMTGFRLLNIDNPHVSSIIIEKW

253 RTSDSRDHTVDMKRPKYTSALTYDGKVMMAEFQSLRRQRIDISRRGNAGDCLANPAVF
 268 SMERLOAPPFPETGLLDGMMTEALMYDAVMVAIASHRASQITVSSLCQCHRHK P
 266 SMERLOAPPFPETGLLDGMMTEALMYDAVMVAIASHRASQITVSSLCQCHRHK P

313 WCGGIDIQALQOQVREGLTQNVQFN EKGRRTNYTIVLEMKRHDGIRKIGYWN EDDKF
 324 WRIGPRFMNLKEARNWGLTGRITFNKTDGLRDFDLDHISLKEEGTKRIGLWNSNSGLN
 322 WRIGPRFMNLKEARNWGLTGRITFNKTNGLRDFDLDHISLKEEGTKRIGLWNSNSGLN

371 VPAATDAQAGGDNSSVQNRITYVTITILEDFYVMLEKKNANOFEGNDRYEGYQVELAAELAK
 384 MTGDRNDRSNNTISLANRTLITVTTILEEPYVMYKSDKPLYGNDRFEGYCLDLIKELSN
 382 MTESQKGPANITISLNSRLIVTTILEEPYVLFKSDKPLYGNDRFEGYCLDLIKELSLT

431 HVGYSTYRLIVSDGKYGARDDPTKAWNGMVGEIVYGRADVAAPLITITVREKVIDFSKF
 444 ILGLYIVDKLVDPDKGYGAON DKGEWNGMVKELIDHRADLAVAPLITITYREKVIDFSKF
 442 ILGLYIIRLVDPDKGYGAODVNGWNGMVRELIDHRADLAVAPLITITYREKVIDFSKF

MSR I

491 FMTLGISIMIKPKQSKPGVFSFLDPLAYETWMCIVFAIGVSVLFLVSRFSPEYENHSE
 503 FMTLGISILYRKPNGTNPGVFSFLNPLSPDINMYVLLACLGVCVLFVIAFRTFYGYWYNF
 502 FMTLGISILYRKPNGTNPGVFSFLNPLSPDINMYVLLACLGVCVLFVIAFRTFYGYWYNF

MSR II

551 EFEEGRDQTTSDQSNFEGIFNSLWFSLGAFMOQGCDSRSLSGRIVGGVWFFTLIIIS
 563 HPCNPDSDVVE NNFTLLNSFWFGVGMALMOGSELMKALSTRIVGGIWWFFTLIIIS
 562 HPCNPDSDVVE NNFTLLNSFWFGVGMALMOGSELMKALSTRIVGGIWWFFTLIIIS

MSR III

611 SYTANLAFLTVRMVSPIDSADDLAKQTEIAYGTLEAGSTKEFTFRSKIAVFEKMMTYM
 620 SYTANLAFLTVRMESPIDSADDLAKQTEIAYGTLEAGSTKEFTFRSKIAVFEKMMTYM
 619 SYTANLAFLTVRMESPIDSADDLAKQTEIAYGTLEAGSTKEFTFRSKIAVFEKMMTYM

671 KSAEPSTVIRTTEGMIRVRKSKGYAYLLESTMEYILHOKPQDTMKVQGNLDSKGYGI
 680 SSRQOSALVNSDEGIQORV LTHDYATLMESTISTEYVTOIN CNLTQIGGLIDSKGYGV
 679 SSRQOSALVNSDEGIQORV LTHDYATLMESTISTEYVTOIN CNLTQIGGLIDSKGYGV

731 ATTKGSALRNPVNLAVKINBOGLDKLNKWNWYDKGECGTGGGSKDKTSALSLSNVNG
 737 GTHTGSPYRDKITITAILOLOEGGLHMKWKWRGN GCPEDSKKE ASVLGVNIGG
 736 GTHTGSPYRDKITITAILOLOEGGLHMKWKWRGN GCPEDSKKE ASVLGVNIGG

791 VFYITIGGGLIAMIVALIEFYKSRSESKRMKGFCLIPQOSINEAIRSTLPRNSGAGAS
 793 IFIVLAAGLVLSVFVAIGFEFLYKSRKNDVQCLSFNATMEELGTSLNKQKLNKRSRTK
 792 IFIVLAAGLVLSVFVAIGFEFLYKSRKNAOLEKRSFSCSAMVEELRMSLKCORRLKHPQPOQ

MSR IV

851 GGGSGENGRRVVSQDFPKSMQSIPCMSSHSGMPLGATGL
 853 GKSSFTSILTCRQRTQRKETVA
 852 LL

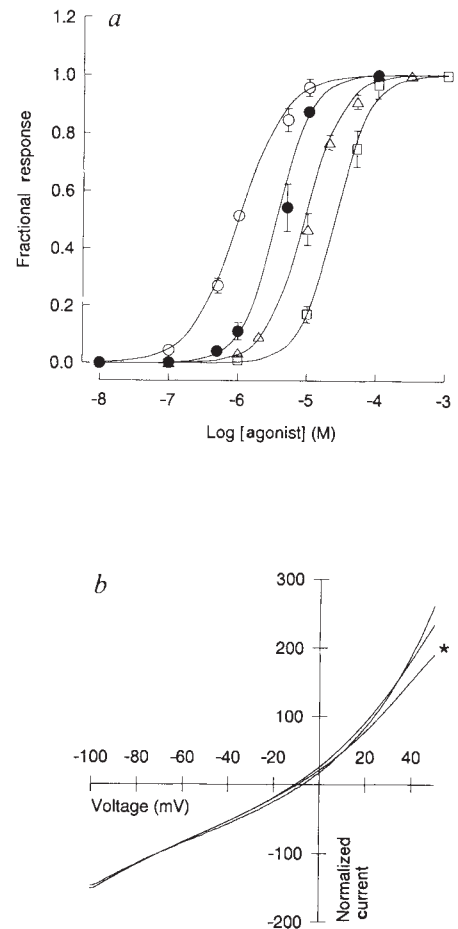


FIG. 3 a, Dose-response curves obtained on GluR6 injected oocytes after conA treatment for kainate ($\circ$), glutamate ($\square$), quisqualate ($\triangle$) and kainate in presence of $10 \mu\text{M}$ CNQX ($\bullet$). All points represent an average $\pm$ s.e.m. 3 to 6 independent measurements. Error bars indicate s.e.m. b, Current-voltage relationship of the homomeric GluR6 receptor evoked by $10 \mu\text{M}$ kainate before and after conA treatment. The asterisk indicates the $I-V$ relationship obtained on conA-treated oocytes in a modified frog Ringer's solution in which the NaCl was substituted with an equimolar concentration of sodium methanesulphonate. The currents were normalized to the individual currents measured at -70 mV (30 nA and $2.3 \mu\text{A}$ in Ringer's solution before and after conA treatment, respectively, and 330 nA in the modified Ringer's solution). METHODS. The $I-V$ relationships were assessed from 2-s voltage ramps from -100 mV to 50 mV in the presence and absence of agonist. Data were collected and analysed using the pClamp program set.

FIG. 1 a, Alignment of the deduced amino-acid sequences of the rat glutamate receptor subunits GluR1, GluR5 and GluR6. The GluR5 clone GluR5-1 lacking a 15-amino-acid insert is used for the alignment⁷. Positions at which the amino acids are identical in at least two proteins are enclosed and shaded. The predicted signal peptide and membrane-spanning regions (MSR) are indicated²². Two models have been suggested for the membrane-spanning regions²⁻⁶, one as indicated here the other in which the region Phe 575-Ser 593 (GluR6 numbers) is predicted to replace MSR I as a membrane-spanning region. Numbers indicate positions in the mature subunits. METHODS. An adult rat cerebellum cDNA library constructed in λ ZAP was screened under conditions of low-stringency hybridization with a GluR5 cDNA fragment⁷. A 3-kb fragment from a cDNA clone encoding part of the GluR6 open reading frame was used to rescreen the library under conditions of high-stringency hybridization. Two clones, RC11 and RC27, possessed sufficient overlap to engineer a cDNA clone encoding the entire open reading frame of the GluR6 protein.

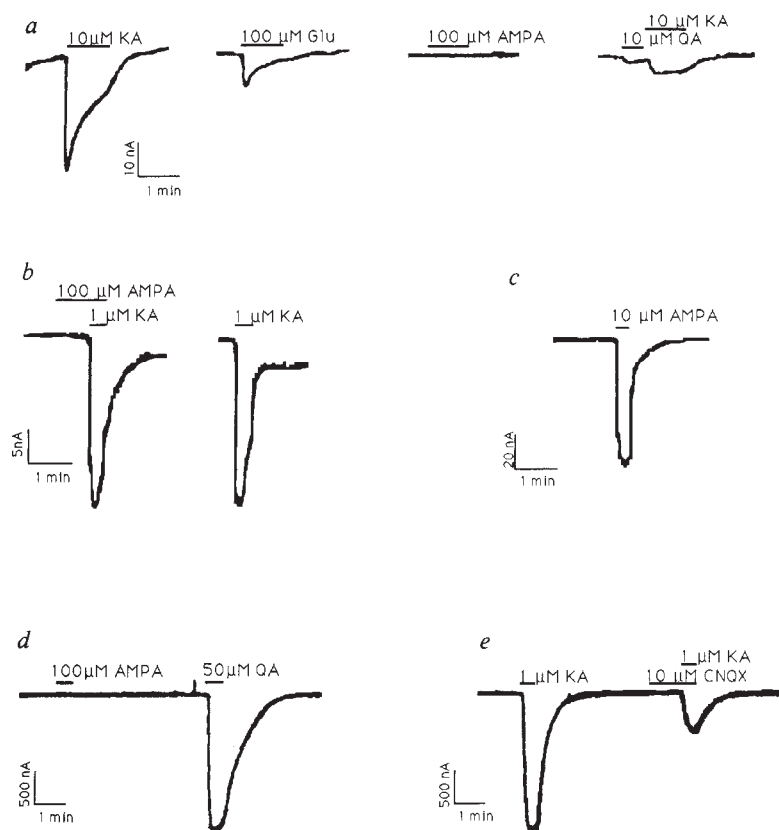


FIG. 2 Agonist evoked rapidly desensitizing currents in *Xenopus* oocytes injected with GluR6 RNA. *a–c*, Current without conA treatment. *d* and *e*, After exposure to 10 μ M conA for 5 min. *a*, Inward currents at a holding potential of -100 mV induced by 10 μ M kainate (KA), 100 μ M glutamate (Glu), or 100 μ M AMPA. Attenuation of the kainate current induced by prior application of quisqualate (QA). All traces are from the same oocyte. *b*, Coapplication of 100 μ M AMPA and 1 μ M kainate followed by a 1 μ M kainate application on the same oocyte. There was 15 min between the two applications to assure full recovery from desensitization. *c*, Current evoked on GluR1 injected oocytes by application of 10 μ M AMPA from the same batch of AMPA as used for the GluR6 experiments. *d*, Application of 100 μ M AMPA or 50 μ M quisqualate. Simultaneous measurement of the current at higher gain did not reveal any AMPA current (<1 nA). *e*, 10 μ M CNQX induced inhibition of 1 μ M kainate response. Bar indicates duration of drug application. Recordings were obtained with different oocytes in *d* and *e*.

METHODS. *Xenopus* oocytes were prepared according to ref. 2 and injected with 10–20 ng *in vitro* synthesized GluR6 RNA. Recordings were performed 4 to 10 days after injection under two electrode voltage clamp using a Dagan 8500 or an Axoclamp-2A amplifier. Both recording and current electrodes were filled with 3.0 M KCl. The recordings were performed in frog Ringer's solution (10 mM HEPES–NaOH, pH 7.2, 115 mM NaCl, 1.8 mM CaCl_2 and 2.5 mM KCl).

been caused either by inhibition of agonist binding or by a direct methanesulphonate block of the channel. The latter effect might be potentiated at positive holding potentials.

The expression pattern of the GluR6 gene was studied by *in situ* hybridization of brain sections from adult mice. The highest numbers of GluR6 transcripts were observed in the olfactory lobe, piriform cortex, dentate gyrus, hippocampus, and in the granular cell layer of the cerebellum. In the hippocampus a gradient in hybridization intensities was observed from rostral to caudal areas, with increased intensity in the CA3 region as compared to the CA1 region (Fig. 4c). The higher number of transcripts in the pyramidal cell layer of CA3 and the granule cell layer of the dentate gyrus correlates with the large amount of [^3H]kainate binding in the stratum lucidum and the commissural/associational terminal field of the dentate gyrus, respectively^{13,18}. Less intense hybridization signals were observed in the caudate putamen, the zona incerta of the thalamus, the inner

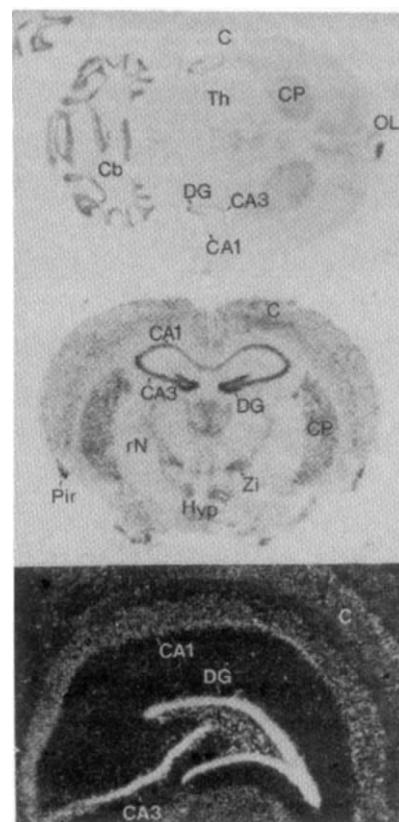


FIG. 4 Autoradiograms showing the localization of GluR6 RNA using *in situ* hybridization histochemistry on adult mouse brain sections. A 1,000-nucleotide ^{35}S -labelled antisense RNA transcript corresponding to the amino terminal part of GluR6 from rat was used as probe. *a*, Horizontal section. *b*, Coronal section. *c*, High-resolution dark-field autoradiogram of hippocampus. C, cortex; CA, Cornu Ammon; Cb, cerebellum; CP, caudate putamen; DG, dentate gyrus; Hyp, hypothalamus; OL, olfactory lobe; Pir, piriform cortex; rN reticular thalamic nucleus; Th, thalamus; Zi, zona incerta.

and outer layers of cortex, several brain stem nuclei as well as the ganglion cell layer of the retina. In general, areas expressing large amounts of transcripts correlate well with areas expressing high-affinity kainate-binding sites.

The properties reported for the homomeric GluR6 receptors have not been described in studies performed on neurons, although glutamate receptors located on afferent C fibres from rat dorsal root ganglions (DRG) also exhibit a fading current in continued presence of kainate and an order of agonist potency similar to what we found for the GluR6 subunit¹⁹. The DRG receptors exhibit an EC_{50} for AMPA of roughly 520 μ M and an EC_{50} for kainate of 13 μ M¹⁹ compared with 100 μ M for kainate in oocytes injected with forebrain mRNA¹⁷. Furthermore, receptors responding to nanomolar concentrations of kainate have been described in hippocampal pyramidal cells isolated from CA3, but they are not found in cells from CA1 (refs 20 and 21), consistent with our observation that GluR6 transcripts are present at a high level in CA3 and a lower level in CA1. CA3 pyramidal cells depolarized reversibly on application of 0.01–1 μ M kainate, whereas 10 μ M kainate apparently killed the neurons²⁰. Thus the pattern of gene expression and the pharmacology of the GluR6 subunit suggest that it corresponds or contributes to the high-affinity receptor for kainate found in the brain. □

Received 25 March; accepted 14 May 1991.

- Collingridge, G. L. & Lester, R. A. J. *pharmac. Rev.* **40**, 143–210 (1989).
- Hollmann, M., O'Shea-Greenfield, A., Rogers, S. & Heinemann, S. *Nature* **342**, 643–648 (1989).
- Boulter, J. et al. *Science* **249**, 1033–1037 (1990).

4. Keinänen, J. *et al. Science* **249**, 556-560 (1990).
5. Nakanishi, N., Shneider, N. A. & Axel, R. *Neuron* **5**, 569-581 (1990).
6. Sakimura, K. *et al. FEBS Lett.* **272**, 73-80 (1990).
7. Bettler, B. *et al. Neuron* **5**, 583-595 (1990).
8. Monaghan, D. T., Bridge, R. J. & Cotman, C. W. *A. Rev. Pharmac. Tox.* **29**, 365-402 (1989).
9. Wada, K. *et al. Nature* **342**, 684-689 (1989).
10. Gregor, P., Mano, I. M., Maoz, I., McKeown, M. & Teichberg, V. I. *Nature* **432**, 689-692 (1989).
11. Meyer, M. L. & Vyklicky, L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1411-1415 (1989).
12. Dawson, T. L., Nicholas, R. A. & Dingleline, R. *Molec. Pharmac.* **38**, 779-784 (1990).
13. Foster, A. C. & Fagg, G. E. *Brain Res. Rev.* **7**, 103-164 (1984).
14. Honore, T. *et al. Science* **241**, 701-703 (1988).
15. Verdoorn, T. A., Kleckner, N. W. & Dingleline, R. *Molec. Pharmac.* **35**, 360-368 (1989).
16. Miledi, R. & Parker, I. *J. Physiol.* **357**, 173-183 (1984).
17. Verdoorn, T. A. & Dingleline, R. *Molec. Pharmac.* **34**, 298-307 (1988).
18. Monaghan, D. T. & Cotman, C. W. *Brain Res.* **252**, 91-100 (1982).
19. Huettner, J. E. *Neuron* **5**, 255-266 (1990).
20. Robinson, J. H. & Deadwyler, S. A. *Brain Res.* **221**, 117-127 (1981).
21. Ben-Ari, Y. & Gho, M. *J. Physiol.* **404**, 365-384 (1988).
22. Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387-395 (1984).

ACKNOWLEDGEMENTS. We thank J. Boulter, R. Papke, R. Hume, M. Hollmann and R. Dingleline for help and advice and J. Boulter for the cerebellum library. This work was supported by The Danish Medical Research Council (J.E.), the Schweizerischer Nationalfond (B.B.), HSFO (I.H.B.), the Fritz B. Burns foundation (S.H.), NINCDS, the Wiengart Foundation (S.H.) and the Human Frontiers Science Program.

A mouse gene related to *Distal-less* shows a restricted expression in the developing forebrain

Melanie Price, Marianne Lemaistre, Michele Pischetola*, Roberto Di Lauro & Denis Duboule

European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, Germany

MANY genes known to be involved in embryogenesis and morphogenesis of the fruitfly *Drosophila melanogaster* encode proteins with a highly conserved region of 60 amino acids called the homeodomain¹. Mammalian counterparts for most of these genes have been identified, including those homologous to the *Drosophila* homeotic genes^{2,3} or to genes such as *evenskipped*⁴, *engrailed*^{5,6} or *caudal*⁷. We have isolated a murine homeobox gene that encodes a homeodomain similar to that encoded by the *Drosophila Distal-less* (*Dll*) gene⁸. *Dll* has a crucial role in *Drosophila* limb morphogenesis, partially specifying pattern along the proximo-distal axis of the limb^{8,9}. The murine counterpart is expressed in a restricted region of the developing brain, within the diencephalon and the adjacent telencephalic regions.

A mouse clone containing sequences homologous to the *Drosophila Dll* gene (*Dlx* clone) was isolated using a DNA probe containing the thyroid transcription factor TTF-1 homeobox sequence¹⁰. This gene maps to mouse chromosome 2 (M.Pr., M. G. Mattei & M.Pi., manuscript in preparation) and its partial genomic structure is shown in Fig. 1a. Sequence analysis of a mouse *Dlx* complementary DNA clone, isolated from a neonatal brain cDNA library, revealed that the region corresponding to the third helix of the *Dlx* homeodomain is interrupted by an intron (between the codons for amino-acid residues 44 and 45, arrow in Fig. 1b). The position of this intron is conserved in the *Drosophila Distal-less* (*Dll*) homeodomain⁸ (Fig. 1b). The *Dlx* homeodomain is 90% identical to that encoded by the *Drosophila* gene, although in the *Drosophila* gene the intron may be as large as 13 kilobases (kb), whereas in the mouse gene it is about 1.4 kb.

Using a *Dlx* RNA probe (*PstI/KpnI* in Fig. 1a), no expression could be detected in adult tissue messenger RNAs, including muscle, lung, thyroid, testes, intestine, thymus, liver, spleen, heart and kidney, or in mRNA from undifferentiated or

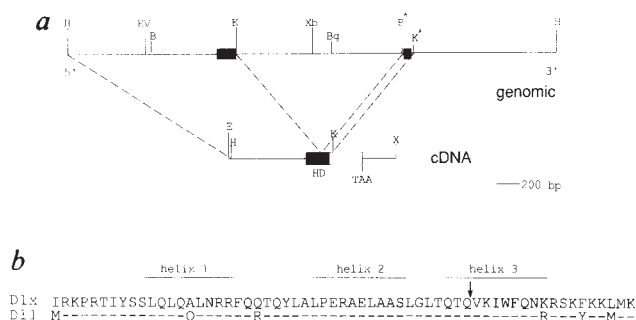


FIG. 1 Partial maps of *Dlx* genomic and cDNA clones. *a*, Partial restriction maps of a 3.8-kb *HindIII* genomic DNA fragment and a 1.4-kb partial cDNA clone which contain sequences hybridizing to the TTF-1 homeobox probe. Black boxes, regions corresponding to the homeodomain (HD); open box, open reading frame in the cDNA, extending from the homeobox to an in-frame stop codon (TAA). Dashed lines indicate equivalent positions in the genomic and cDNA clones. A second intron lies 5' to the homeobox as indicated by the absence of EV and B sites in the cDNA. The exact location of this intron has not been determined. H, *HindIII*; EV, *EcoRV*; B, *BamHI*; K, *KpnI*; Xb, *XbaI*; Bg, *BglII*; P, *PstI*; X, *XhoI*. The asterisks above the *PstI* and *KpnI* sites in the genomic map indicate the DNA fragment which was subcloned into a Bluescript vector and used as a probe in expression analysis. The probe spans from amino acid 42 in the third helix to 24 amino acids downstream of HD. *b*, Deduced amino-acid sequence (single-letter code) of the *Dlx* homeodomain and its relationship to the *Drosophila Dll* homeodomain. The dashes indicate positions where amino acids are shared with *Dlx*. The positions of the α helices of the homeodomain are from ref. 1. Both homeodomains are encoded by a gene with an intron between codon 44 and 45 (arrowhead).

METHODS. A mouse genomic library (5×10^6 plaque-forming units was screened with a randomly primed probe containing the sequence of the rat TTF1 homeodomain¹⁰. Filters were washed at low stringency ($2 \times$ SSC, 50°C) and clones containing related sequences were identified by sequencing. The *PstI/KpnI* probe was used to screen a neonatal mouse brain library (Stratagene) and the homeobox positive clones were sequenced. The *Dlx* sequence has been submitted to the EMBL data bank (Accession number X57552).

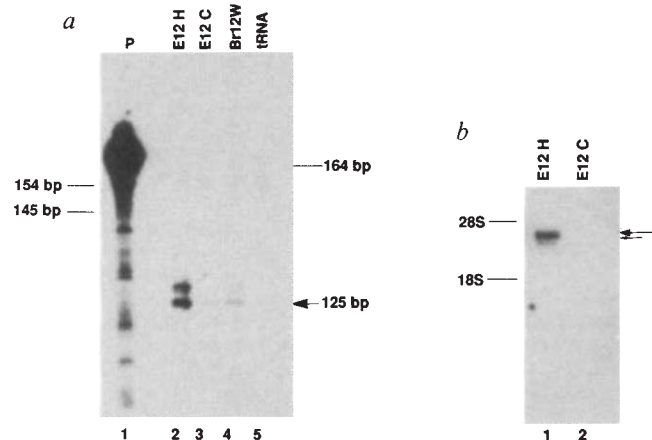


FIG. 2 Expression of *Dlx* RNA in mouse brain. *a*, Total RNA (6 µg) isolated from the heads (E12H) or carcasses (E12C) of day-12.5-p.c. mouse embryos, 12-week-old mouse brain (Br12W) or transfer RNA (tRNA) was hybridized with 50,000 c.p.m. antisense *PstI/KpnI* riboprobe transcribed using the T3 promoter in a Bluescript vector. The probe length is 164 base pairs (bp) (P, lane 1). A protected fragment of the expected size (125 bp, arrowhead) is found in embryonic head RNA (E12H, lane 2) and adult brain RNA (Br12W, lane 4) but not in RNA from embryonic carcass (E12C, lane 3) or tRNA (lane 5). Two bands are always observed in RNase mapping using this probe. *b*, Total E12H and E12C RNA (50 µg) was fractionated on a formaldehyde-agarose gel and hybridized with ^{32}P -labelled *PstI/KpnI* probe. An abundant transcript is observed (arrowhead) in E12H RNA (lane 1), just below the 28S ribosomal RNA, and a less abundant transcript is observed below this. METHODS. Northern blots and RNase protection experiments were carried out according to refs 20 and 21, respectively.

* Present address: Stazione Zoologica 'A. Dohrn', Villa Comunale, Napoli, Italy.

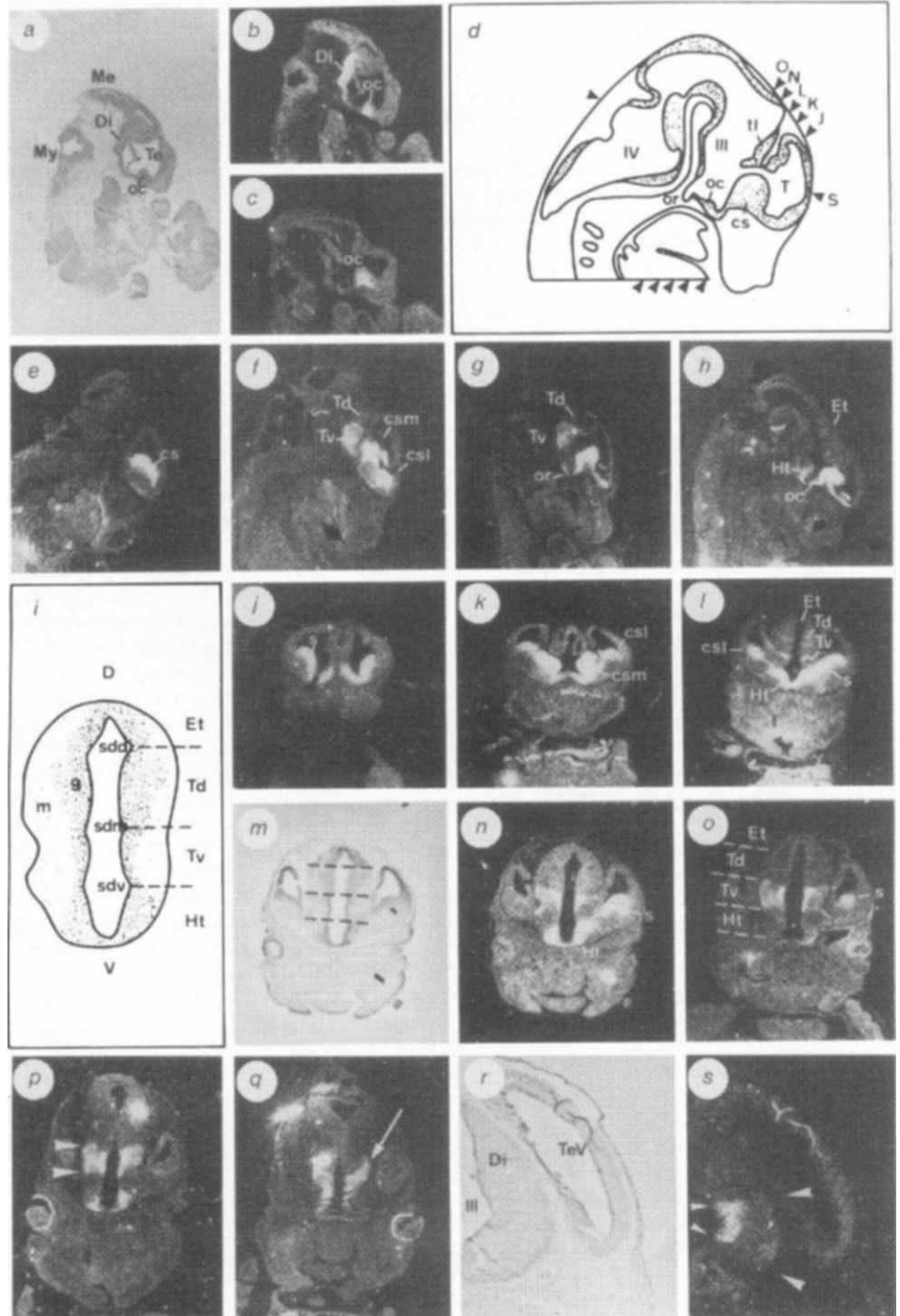
differentiated P19 and F9 embryonic stem cells (data not shown). *Dlx* mRNA was found in RNA prepared from heads of day-12.5 post-coitum (p.c.) embryos (Fig. 2a). A reduced expression was also detected in RNA prepared from 3-month-old murine brains (Fig. 2a). Northern blot analysis revealed an abundant transcript of about 4.5 kb and a less abundant transcript of slightly smaller size (Fig. 2b, arrows).

We analysed in detail the *Dlx* expression pattern in the fetal brain by *in situ* hybridization. Sagittal, transverse and frontal sections of mouse embryos from days 8.5 to 15 p.c. were analysed using an antisense RNA probe. The earliest time that *Dlx* transcripts were detected was at day 10 p.c., when they were seen in the region of the optic chiasma (Fig. 3a-c). The transcripts

also appear in the walls of the diencephalon (Fig. 3b). At day 12 p.c., the *Dlx* expression pattern in the brain is established and can be seen best by looking at all three planes of section. At this stage, the floors of the telencephalic vesicles are thicker owing to the high mitotic activity of the germinative zone. The striatum (corpus striatum) will differentiate from this region, giving rise later on to the central nuclei. The striatum is believed to be of both diencephalic (median striatum or paleostriatum) and telencephalic (lateral striatum or neostriatum) origin. These areas later give rise to structures such as the globus pallidum, nucleus caudate and putamen. *Dlx* is strongly expressed in the striatum germinal zone (Fig. 3e-l), with the region in which the transcripts are present extending widely along the

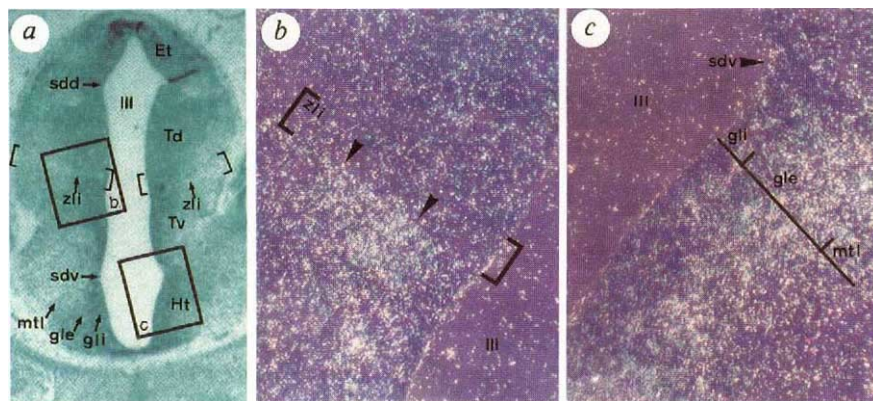
FIG. 3 Expression pattern of *Dlx* during forebrain development as observed by *in situ* hybridization. a-c, Sagittal sections of a day-10-p.c. fetus hybridized with a *Dlx*-specific antisense RNA probe. The bright-field picture (a) is a more paramedian section and corresponds to the dark-field picture shown in b, c. A more median plane of section from the same embryo. d-h, Sagittal serial sections through a day-12.5-p.c. fetal head. The schematic diagram in d corresponds roughly to the section shown in h and is rather central (note the continuity between the different vesicles). Sections shown in g, f and e are progressively more lateral, respectively. Section shown in e is a very lateral section showing specific *Dlx* expression in the corpus striatum (cs), whereas more median sections show expression in thalamic (Tv) and hypothalamic (Ht) regions as well. i, Scheme of a frontal section through the 12.5-day-p.c. developing diencephalon. The scheme is enlarged from m, n. The four Herrick's zones and the different sulci are indicated. j-o, Serial frontal sections of a day-12.5-p.c. fetal head. The various planes of section are indicated by arrowheads in d, j, very frontal section, with sectioning towards the dorsal part of the head. m, Bright-field view corresponding to n; dashed lines in m and o indicate the separations between the various parts of the diencephalon described for i, f, q. Two sections from another fetus at the same age; p is more frontal than q (p is at about the same position as o). Arrowheads in p point to the one-segment-like expression of *Dlx* in the ventral thalamus, and the white arrow in q shows the extension of the *Dlx* expression region in more lateral parts, probably in the area presumed to correspond to the future corpus geniculatum ventrale. r, s, Bright-field and dark-field picture of a transversal section roughly at the level shown by the arrowhead marked s in d. It shows the restriction of *Dlx* expression to a central region along this axis (arrowheads in s). My, myelencephalon; Me, mesencephalon; Di, diencephalon; Te, telencephalon; oc, optic chiasma; IV, fourth ventricle; III, third (diencephalic) ventricle; or, optic recess; T, telencephalic vesicles; cs, corpus striatum; tl, thalamus; Td, dorsal thalamus; Tv, ventral thalamus; csm, corpus striatum mediale; csl, corpus striatum laterale; Et, epithalamus; Ht, hypothalamus; s, striatum; m, mantle layer; g, germinative layer; D, dorsal; V, ventral; sdd, sulcus diencephalicus dorsale; sdm, sulcus diencephalicus mediale; sdv, sulcus diencephalicus ventrale.

METHODS. Mouse embryos and fetuses were obtained from normal matings



between (C57BL/6 × SJL) F₁ mice. Day 0.5 p.c. was assumed to begin at the middle of the day of vaginal plugging. *In situ* hybridizations were carried out as described in ref. 22, except that the prehybridization step was omitted.

FIG. 4 *Dlx* expression in the day-12.5-p.c. diencephalon. *a*, An enlargement of Fig. 3*m* which shows the division of the diencephalon into its four regions. *b*, *c*, Enlargements of subregions indicated by the rectangles in *a*. *b*, Presence of the zona limitans intrathalamica (square brackets in both *a* and *b*) is shown. The hybridization grains are found immediately below (ventral to) this zone but not inside or dorsally, as indicated by arrowheads. *c*, Enlargement of the anterior hypothalamic area (at the level of the optic chiasma) with the grains found in the external germinative layer and in the mantle layer but not in the ependymal layer. zli, Zona limitans intrathalamica; mtl, mantle layer; gli, internal germinative (ependymal) layer; gle, external germinative layer. All other abbreviations are as in Fig. 3. Methods as for Fig. 3.



anteroposterior (see, for example, Fig. 3*e*) and mediolateral (Fig. 3*l*) axes. The transcript is thus expressed in very frontal (Fig. 3*f*) and lateral (Fig. 3*k*, *l*) areas as well as very posteriorly (Fig. 3*o*). The two lateral expression areas are separated in frontal regions (Fig. 3*j*) but fuse in a more posteromedian zone, where the telencephalon and diencephalon are contiguous, roughly at the level of the optic recess, ventrally (Fig. 3*l*). The diencephalon of mouse embryos from days 12–13 p.c. is at the stage of the differentiation of the germinal mantle and marginal layers¹¹ (Fig. 4*a*). It is divided into Herrick's four zones—epithalamus, dorsal thalamus, ventral thalamus and hypothalamus—by the ventricular sulci (the sulcus diencephalic dorsale, mediale and ventrale)¹² (Fig. 3*l*). These sulci correspond to morphological landmarks separating zones that are phylogenetically and functionally independent. This discontinuous appearance is illustrated by the separation between the dorsal and ventral thalami (the equivalent of the thalamic and subthalamic areas in human) which occurs at the level of the sulcus medians (Figs 3*l*, 4*a*) and is reinforced by the presence of the zona limitans intrathalamica (Fig. 4*a*, zli), a layer of low-density cell population which lies orthogonal to the major ventricular axis¹² (see, for example, Fig. 3*p*, arrowheads, and Fig. 4*b*). *Dlx* is expressed immediately below the zona limitans intrathalamica whose cells do not contain *Dlx* transcripts (Fig. 4*b*). The 'dorsal' (ontogenetically posterior because the diencephalon rotates by 90° during development) limit of the *Dlx* expression region (upper arrowhead in Fig. 3*p*) is therefore orthogonal to the third ventricle (Fig. 4, III) and extends dorsally in more lateral areas (Fig. 3*q*, arrow), probably corresponding to the future corpus geniculatum ventrale¹². Anteriorly, *Dlx* is expressed in the entire ventral thalamus according to a distinctive segment-like pattern (Fig. 3*q*, *p*). More posteriorly, however, the transcript domain extends ventrally within the hypothalamus (Fig. 3*q*) where *Dlx* transcripts are present in central and lateral areas (Fig. 4*c*), only in the anterior part, roughly to the level of the optic chiasma (Fig. 3*p*, *q*). The hypothalamus has no *Dlx* transcripts posterior to the optic chiasma, around Rathke's pocket and the region of the future neurohypophysis (not shown). *Dlx* expression in both the ventral thalamus and the hypothalamus is observed in para-ventricular and more lateral areas (Fig. 4), but is not detected in the internal germinal layer (gli). Only the external germinal layer (gle) and the mantle layer show hybridization grains (Fig. 4). The apparent differences in grain densities between the germinative and mantle zones (for example, Fig. 3*p* or *s*) is due to differences in cell densities rather than to differential gene expression. The same pattern of restricted *Dlx* expression is observed in transverse sections, which indicates that *Dlx* is not expressed either very anteriorly or very posteriorly in the ventral thalamus, but rather in a central area, vertical to the optic chiasma (Fig. 3*r*, *s*).

Thus the regions of expression of *Dlx* in the developing forebrain do not seem to correspond to either a cell layer or to prospective zones for particular central, thalamic or

hypothalamic nuclei. *Dlx* transcripts are instead present in several layers of areas presumed to correspond to different nuclei such as the median and lateral corpus striatum, the zona incerta, nucleus reticularis, corpus geniculatum ventrale or the preoptic nuclei, as well as to other nuclei in the anterior hypothalamic region.

The class I homeobox-containing genes (*Hox*) are expressed in the developing central nervous system^{13,14}, but never anterior to the myelencephalon. At least one of these genes (*Hox-2.9*) is expressed according to a one-segment-like pattern^{15–17} related to the presence of morphologically distinct rhombomeric bulges¹⁸ (neuromeres). The developing vertebrate diencephalon also displays neuromeric patterns which subdivide the forebrain into functionally distinct areas (ref. 19 and refs therein). The *Dlx* expression domain is observed within morphological landmarks, from the most anterior part of the diencephalon until the zona limitans (sulcus median), which corresponds to the interparencephalic neuromeric boundary described in ref. 19. The *Dlx* gene product could therefore be involved in the early development of diencephalic subdivisions.

It is difficult to find a parallel between the pattern of *Dlx* expression in mammals and the function of *Dll* in insects. In mammals, however, *Dlx* is expressed around the optic chiasma and in regions presumed to correspond to the future preoptic nuclei and ventral corpus geniculatum; all of these areas are involved in the reception and the transmission of the optic signals from the optic nerves to the proper cortical areas. In *Drosophila*, *Dll* is expressed in the optic centre of the developing larval brain (S. Cohen, personal communication). Also, *Dlx* expression is seen in the mesoderm of the first branchial arches (not shown). So was an ancestral *dll*-like gene recruited during evolution so that it now achieves partially different functions in insects and mammals? Indeed, the *Dlx* gene is expressed in regions of the rodent brain that are believed to be phylogenetically old (the palaeencephalon) and that develop earlier during ontogeny and represent most, if not all, of the forebrain of less evolved vertebrates. □

Received 6 February; accepted 3 May 1991.

1. Scott, M. P., Tamkun, J. W. & Hartzel, G. U. *Biochim. biophys. Acta* **969**, 25–48 (1989).
2. Duboule, D. & Dollé, P. *EMBO J.* **8**, 1497–1506 (1989).
3. Graham, A., Papalopulu, N. & Krumlauf, R. *Cell* **57**, 367–374 (1989).
4. Bastian, H. & Gruss, P. *EMBO J.* **9**, 1839–1845 (1990).
5. Joyner, A. & Martin, G. *Genes Dev.* **1**, 29–36 (1987).
6. Davies, C. et al. *Genes Dev.* **2**, 1736–1743 (1988).
7. Duprey, P. et al. *Genes Dev.* **2**, 1647–1655 (1988).
8. Cohen, S. M., Brunner, G., Kümmel, F., Jürgens, G. & Jäckle, A. *Nature* **338**, 432–434 (1989).
9. Cohen, S. M. & Jürgens, G. *EMBO J.* **8**, 2045–2056 (1989).
10. Guazzi, S. et al. *EMBO J.* **9**, 3631–3639 (1990).
11. Rugh, R. *The Mouse, Its Reproduction and Development* (Burgess, Minneapolis, 1968).
12. Niimi, K. et al. *Tohoku exp. Med.* **8**, 203–238 (1962).
13. Gaunt, S. J., Sharpe, P. T. & Duboule, D. *Development* **104**, 71–82 (1988).
14. Holland, P. W. H. & Hogan, B. L. M. *Genes Dev.* **2**, 773–782 (1988).
15. Wilkinson, D. G. et al. *Nature* **341**, 405–409 (1989).
16. Murphy, P., Davidson, D. & Hill, R. E. *Nature* **341**, 156–159 (1989).
17. Wilkinson, D. G. & Krumlauf, R. *Trends Neurosci.* **8**, 335–339 (1990).
18. Lumsden, A. *Trends Neurosci.* **8**, 335–339 (1990).

19. Puelles, L. *et al.* *J. comp. Neurol.* **266**, 247–268 (1987).
20. Church, N. G. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991–1995 (1984).
21. Glover, D. M. *DNA Cloning: A Practical Approach* (IRL, Oxford, 1989).
22. Dollé, P. & Duboule, D. *EMBO J.* **8**, 1507–1515 (1989).

ACKNOWLEDGEMENTS. We thank P. Gruss, U. Deutsch and M. de Felice for the mouse genomic library and RNA samples; J.-P. Changeux, D. Lazzaro, A. Prochiantz, C. Stern and D. Wilkinson for comments and suggestions on the manuscript; and the staff of the EMBL photolaboratory for preparing the figures.

Bell-shaped calcium-response curves of $\text{Ins}(1,4,5)\text{P}_3$ - and calcium-gated channels from endoplasmic reticulum of cerebellum

Ilya Bezprozvanny*†, James Watras*
& Barbara E. Ehrlich*‡§

Departments of Medicine* and Physiology‡, University of Connecticut Health Center, Farmington, Connecticut 06032, USA

RELEASE of calcium from intracellular stores occurs by two pathways, an inositol 1,4,5-trisphosphate (InsP_3)-gated channel^{1–3} and a calcium-gated channel (ryanodine receptor)^{4–6}. Using specific antibodies, both receptors were found in Purkinje cells of cerebellum^{7,8}. We have now compared the functional properties of the channels corresponding to the two receptors by incorporating endoplasmic reticulum vesicles from canine cerebellum into planar bilayers. InsP_3 -gated channels were observed most frequently. Another channel type was activated by adenine nucleotides or caffeine, inhibited by ruthenium red, and modified by ryanodine, characteristics of the ryanodine receptor/channel⁶. The open probability of both channel types displayed a bell-shaped curve for dependence on calcium. For the InsP_3 -gated channel, the maximum probability of opening occurred at 0.2 μM free calcium, with sharp decreases on either side of the maximum. Maximum activity for the ryanodine receptor/channel was maintained between 1 and 100 μM calcium. Thus, within the physiological range of cytoplasmic calcium, the InsP_3 -gated channel itself allows positive feedback and then negative feedback for calcium release, whereas the ryanodine receptor/channel behaves solely as a calcium-activated channel. The existence in the same cell of two channels with different responses to calcium and different ligand sensitivities provides a basis for complex patterns of intracellular calcium regulation.

Addition of InsP_3 to endoplasmic reticulum vesicles isolated from canine cerebellum induces release of roughly 30% of the accumulated calcium^{9,10}. This magnitude of calcium release is comparable to that reported for InsP_3 -induced calcium release from sarcoplasmic reticulum vesicles from aortic smooth muscle^{3,11} and suggests that 30% of the vesicles have at least one InsP_3 -gated channel. After fusion of these vesicles with planar lipid bilayers, channel activity was not observed in the absence of InsP_3 (Fig. 1, top two traces). Although addition of adenine nucleotides did not activate the InsP_3 -gated channel, AMP-PCP, a poorly hydrolysable analogue of ATP, was added to increase channel activity as described previously^{3,12,13}. Similarly, addition of calcium in the absence of InsP_3 did not activate the InsP_3 -gated channel (Fig. 1, second pair of traces shows current in the presence of both AMP-PCP and calcium; when either compound was added alone, activity was not induced (data not shown)). Addition of InsP_3 induced channel

openings in a concentration-dependent manner (Fig. 1, bottom four traces, ref. 9). Analysis of the current amplitudes reveals that the single channel currents are complex with four conductance levels, each a multiple of 20 pS, with openings to level three predominating (described in ref. 9). These InsP_3 -gated single channel currents are not affected by 2 μM ruthenium red or 5 μM ryanodine (data not shown) and are inhibited by 10 μM heparin³.

InsP_3 receptors are regulated by calcium. InsP_3 binding to the InsP_3 receptor is blocked by calcium with half-maximal

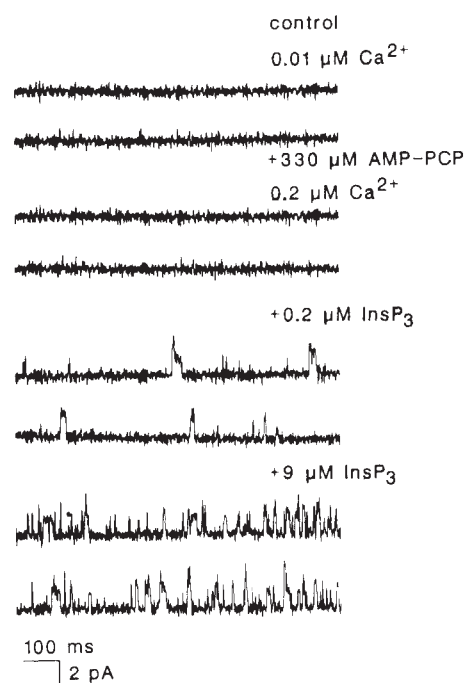


FIG. 1. InsP_3 -gated calcium channels from endoplasmic reticulum of cerebellum in planar lipid bilayers. In the absence of InsP_3 , channel activity was not observed. Addition of AMP-PCP (330 μM) and calcium (0.2 μM free calcium) also did not activate the channel. Subsequent addition of 0.2 μM InsP_3 to the cytoplasmic side induced channel activity with single channel currents represented by four conductance levels, each a multiple of 20 pS (ref. 9). A second addition of InsP_3 , raising the total concentration to 9 μM , further increased channel activity. Channel openings indicating divalent cation movement from *trans* to *cis* side are shown as upward deflections from zero current.

METHODS. Endoplasmic reticulum vesicles derived from canine cerebellum were made as described previously^{9–11}. Bilayers made from phosphatidyl-ethanolamine and phosphatidylserine (75%:25%; Avanti Polar Lipids), were formed by painting a solution of lipid in decane across a 100- μm hole in a Teflon sheet that bisected a Lucite chamber. Channel incorporation into the bilayer was followed as described previously³ with minor modifications. Briefly, vesicles (~5 μg protein) were added to the *cis* chamber which contained 800 mM KCl, 250 mM HEPES-Tris, 10 mM CaCl_2 , pH 7.35. The *trans* chamber contained 53 mM $\text{Ca}(\text{OH})_2$, 250 mM HEPES, pH 7.35, and was held at virtual ground. Solutions in both chambers were stirred until potassium and/or chloride channels appeared as indicators of fusion of vesicles to the bilayer. Then, the *cis* chamber was perfused with 10 volumes of HEPES-Tris solution (250 mM HEPES-Tris, pH 7.35, 1 mM HEDTA) leaving calcium on the *trans* side of the bilayer as the only small ion present in the system. All measurements were made with a transmembrane potential of 0 mV. The *cis* side corresponds to the cytoplasmic side of the reticular membrane and all compounds (AMP-PCP, ATP, CaCl_2 , InsP_3) were added to the *cis* side. Vesicle fusion occurred in virtually all attempts, but InsP_3 -gated channels were observed after perfusion in 35% of our experiments. Note that the channel open times are short (mean ~3 ms). In addition, the greatest open probability we observed in any experiment was 15% using optimal conditions. Data were stored on a digital tape recorder. For computer analysis using pClamp 5.5, data were filtered to 500 Hz by an eight-pole Bessel filter and digitized at 2 kHz. Data shown are examples selected from at least four similar experiments unless otherwise stated.

* Permanent address: Institute of Cytology, Academy of Sciences of the USSR, Leningrad 194064, USSR.

§ To whom correspondence should be addressed.

inhibition at 0.3 μM calcium^{14,15}, and is probably mediated by protein(s) distinct from the InsP_3 receptor, for example, 'calmodin'¹⁶. In permeabilized smooth muscle cells, InsP_3 -induced calcium release depends on cytoplasmic calcium in a biphasic manner with maximal release occurring at a free calcium concentration of 0.3 μM (ref. 13). Over the concentration range 0.01 to 0.3 μM , however, calcium activates InsP_3 -induced calcium release¹³. The mechanism for the activation by calcium below 0.3 μM could occur by a calcium dependence of InsP_3 metabolism or by a direct effect on the channel's gating mechanism. Use of permeabilized cells did not allow discrimination between these possibilities¹³.

To test whether there is a direct effect of calcium on the gating of the channel, we monitored the effect of cytoplasmic calcium on channel activity in planar bilayers. The open probability of

the channel increased as the free calcium concentration was elevated from 0.01 μM to 0.25 μM and decreased at concentrations above 0.25 μM (Fig. 2a). It was possible to go repeatedly from activating levels of calcium to inactivating levels of calcium with the expected changes in channel activity. Resulting data-points from the average of four experiments show that the bell-shaped calcium-dependence curve of these channels decreases sharply on both sides of the maximum (Fig. 2b) and that the entire curve falls within the physiological range of calcium concentrations.

Addition of AMP-PCP without InsP_3 did not activate the InsP_3 -gated channels (Fig. 1), but in some experiments adenine nucleotides activated another type of channel very different from the InsP_3 -gated channels in kinetics and pharmacology (Fig. 3a). This other channel had a conductance of 50 pS (subcon-

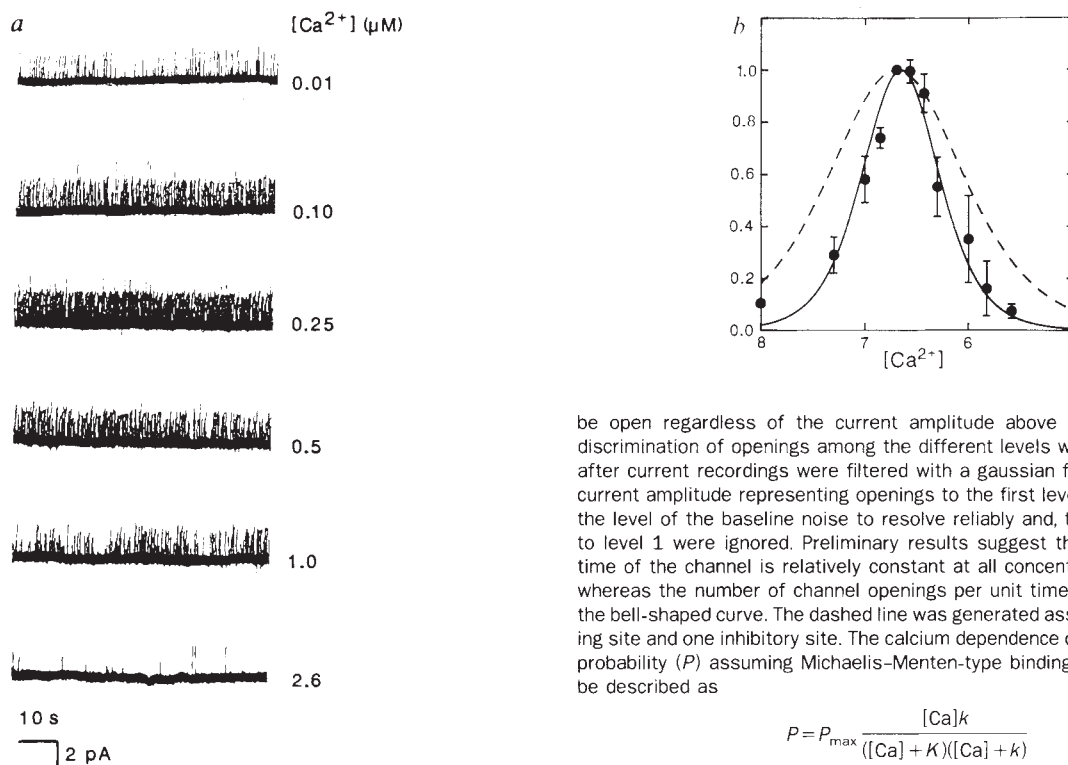


FIG. 2 Calcium-dependence of the InsP_3 -gated channels in planar lipid bilayers. *a*, Channel activity was monitored in the presence of 2 μM InsP_3 , 330 μM AMP-PCP, and at a transmembrane potential of 0 mV. To change the free calcium concentration in the range from 0.01 μM to 5 μM , aliquots of a calibrated 20 mM CaCl_2 solution were added to a mixture of 1 mM EGTA and 1 mM HEEDTA. The free calcium concentration, indicated in μM at the right side of each record, was calculated using binding constants for calcium-EGTA and calcium-HEEDTA complexes of 0.12 μM and 2.2 μM at pH 7.35, respectively³⁴. The effect of calcium was reversible. Note that the time scale is compressed. *b*, Open probability of the InsP_3 -gated channel versus the free calcium concentration ($[\text{Ca}]$). Using a threshold level of 1 pA, the open probability of the channel was calculated from recordings at least 2.5 min long at each calcium concentration. To compare four different experiments using vesicles from two different preparations, values were normalized to the maximum channel activity observed in each experiment, which never exceeded 15% assuming one channel in the bilayer. Thus, the ordinate represents channel activity normalized to fall between 0 and 1. Values plotted are mean $\pm$ s.e.m., $n=4$. To quantitate the results we had to make several assumptions because the channel opens to four different conductance levels⁹ and because channel openings are very fast. These channels primarily open to the 60 pS level (corresponding to level 3), which is a current of 1.8 pA at 0 mV (ref. 9). During data processing we set the threshold for channel opening to 1 pA, a level that avoids the baseline noise, but includes openings to levels 2, 3 and 4. The channel was considered to

be open regardless of the current amplitude above the threshold, and discrimination of openings among the different levels was not made. Even after current recordings were filtered with a gaussian filter to 500 Hz, the current amplitude representing openings to the first level was too close to the level of the baseline noise to resolve reliably and, therefore, openings to level 1 were ignored. Preliminary results suggest that the mean open time of the channel is relatively constant at all concentrations of calcium, whereas the number of channel openings per unit time changes to match the bell-shaped curve. The dashed line was generated assuming one activating site and one inhibitory site. The calcium dependence of the channel open probability (P) assuming Michaelis-Menten-type binding to both sites can be described as

$$P = P_{\max} \frac{[\text{Ca}]k}{([\text{Ca}] + K)([\text{Ca}] + k)} \quad (1)$$

where $P_{\max}$ is the maximum open probability observed, K and k are the binding constants for the activation and inhibitory sites, respectively, and $[\text{Ca}]$ is the free calcium concentration on the cytoplasmic side of the channel. The solid line was generated by two alternative modifications in the model. Modification 1 assumes that calcium binds cooperatively with a Hill coefficient n . Channel open probability (P_1) will be

$$P_1 = P_{\max} \frac{[\text{Ca}]^n k^n}{([\text{Ca}]^n + K^n)([\text{Ca}]^n + k^n)} \quad (2)$$

The best fit to our data can be obtained using $n=1.8$, $K=k=0.2 \mu\text{M}$. This model suggests that calcium binds cooperatively to at least two sites to open the channel and to two other sites to close the channel and that the affinity for calcium to the activation and inhibitory sites is equivalent. With modification 2 of the initial model, we assumed that every InsP_3 channel complex is composed of several independent subunits. If each subunit has one activating and one inhibitory binding site, and all subunits equally affect the channel open probability, then channel open probability (P_2), assuming m equal number of subunits, would be:

$$P_2 = P^m \quad (3)$$

where P is the function from equation (1). Surprisingly, this model also gave a good fit to the data using $m=2.7$, $K=k=0.2 \mu\text{M}$. Note that the affinity for calcium was the same as above, but with this model at least three independent sites are required for activation and three independent sites for inhibition. The curves generated by the two different modifications, however, are indistinguishable.

ductance levels also were observed; data not shown). The channel also was activated by 5 mM caffeine and was completely blocked by 2 μ M ruthenium red (data not shown), suggesting that cerebellum contains functional ryanodine receptors. In support of this hypothesis, addition of 1 μ M ryanodine reduced the channel conductance and channel gating was slowed (Fig. 3a, third pair of traces). When binding assays were done using the same preparation of endoplasmic reticulum vesicles, both InsP_3 and ryanodine binding sites were found (B_{max} , 35 pmol mg^{-1} and 0.18 pmol mg^{-1} , respectively). In the bilayer, InsP_3 -gated channels were more prevalent than the other channel type (ratio >20:1).

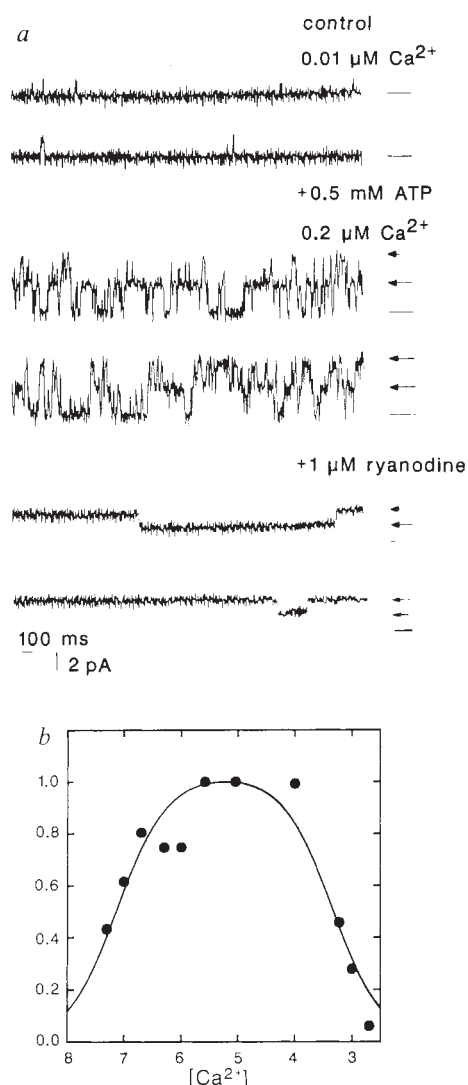


FIG. 3 A second type of channel in cerebellum reticulum membranes. *a*, Experimental conditions were the same as in Fig. 1. In this case, however, after addition of 500 μM ATP and increasing free calcium to 0.2 μM , a channel, different from the InsP_3 -gated channel, was activated. The second pair of traces shows current fluctuations from two channels in the presence of both ATP and calcium; activity could be induced by addition of either compound alone. Addition of ryanodine (1 μM) altered the conductance and the rate of channel opening and closing (third pair of traces). Zero current is indicated by lines in the right margin and arrows indicate current amplitude for one and two open channels. The conductance of the second type of channel was estimated to be 50 pS from current amplitude measurements between -20 and 20 mV. *b*, Open probability of the calcium-gated channels as a function of the free calcium concentration ($[\text{Ca}]$). The open probability of these channels was normalized to the maximum value and, thus, the ordinate represents channel activity normalized to fall between 0 and 1. In the physiological range of cytoplasmic calcium these channels behave solely as calcium-activated channels.

The calcium dependence of the second channel type was biphasic with maximum activity maintained between 1 μM and 100 μM free calcium (Fig. 3b). A calcium-induced decrease in channel activity occurs at calcium concentrations above values normally found within cells.

Although the conductance of the adenine nucleotide and calcium-activated channel (50 pS) is different from the conductance of the ryanodine receptor from skeletal or cardiac ventricular muscle (120 pS and 100 pS, respectively^{3,6,17}), the pharmacology of the channels was indistinguishable. Additional variations in the conductance of the calcium-release channel were found when comparing different regions of the heart (100 pS in left ventricular free wall and 31 pS in interventricular septum¹⁸). These differences in conductance may reflect structural differences in channel proteins. Indeed, isoforms of the ryanodine receptor were identified within skeletal muscle tissue using ryanodine receptor-specific antibodies¹⁹.

Models based on biochemical and structural analyses suggest that the InsP_3 -gated channel is a tetramer^{20–24}. Calcium release and single-channel data suggest that there are functional interactions among the subunits^{9,25}. Nonetheless, we started by constructing a simple model with just two independent binding sites for calcium to describe the calcium dependence, an activation site to open the channel and an inhibitory site to close the channel. Curves generated with this model are bell-shaped but are not sharp enough to fit the experimental data (Fig. 2b, dashed line).

We modified the model to make the theoretical curve sharper. First, we assumed that calcium binds cooperatively at both activation and inhibitory sites. Alternatively, we assumed that every InsP_3 channel complex is composed of several independent subunits. Surprisingly, both models generated good fits to the data (Fig. 2, solid line represents independent model; see legend for corresponding equations and parameters). The curve generated by the cooperativity model is virtually indistinguishable; we can not discriminate between these models. In either case, both models require the existence of a multimeric structure of the InsP_3 channel complex with functional interactions among the subunits^{9,20–24}.

Both InsP_3 and ryanodine receptors exist in the same cell of brain^{7,8}, liver²⁶ and smooth muscle (ref. 27 and our unpublished observations). We have shown that two functional calcium-release channels are present in cerebellum. Even though the ryanodine receptor is relatively low in abundance, its activation could still contribute to the alteration of cytoplasmic calcium levels as predicted in some models for calcium oscillations^{28,29}, especially if the ryanodine receptor controls release of calcium from a separate pool of calcium that is insensitive to InsP_3 .

That the InsP_3 -gated channel has a biphasic dependence on calcium at concentrations found in cells demonstrates another biologically important level of regulation of this channel. Thus, the InsP_3 -gated channel can act as a calcium-gated channel in the presence of a fixed concentration of InsP_3 if the cytoplasmic calcium is less than 0.3 μM . Calcium-induced calcium release has often been an indicator for the presence of the ryanodine receptor; our observations makes this postulate less absolute. Our data also provide a mechanism of positive feedback that can be added to models of calcium oscillators that assume the presence of only InsP_3 -gated calcium-release channels^{30,31}. Calcium stimulation of InsP_3 production (see, for example, ref. 32) can be complementary to the calcium-induced amplification of channel activity in cells.

The sharp decrease in activity of the InsP_3 -gated channel on both sides of the optimal calcium concentration allows us to make some conclusions about the physical structure of the InsP_3 -gated channel complex. Depending on the model used, at least two, and possibly up to four, calcium ions must bind to the InsP_3 -gated channel complex to alter channel open probability. This value is consistent with biochemical estimates^{20–24} and electron microscopy^{22,23}, which predict that the InsP_3 -gated

channel is a tetramer. The association of four receptors is also supported by the observation of four conductance levels with channels from native membranes⁹ and from purified, reconstituted receptor³³.

Note added in proof. While this paper was under review a paper has been published³⁵ reporting that InsP_3 -induced calcium release from brain synaptosomes also displays a bell-shaped dependence on calcium. $\square$

Received 7 February; accepted 10 May 1991.

1. Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67–68 (1983).
2. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315–321 (1984).
3. Ehrlich, B. E. & Watras, J. *Nature* **336**, 583–586 (1988).
4. Endo, M. *Physiol. Rev.* **57**, 71–108 (1977).
5. Meissner, G., Darling, E. & Eveleth, J. *Biochemistry* **25**, 236–244 (1986).
6. Smith, J. S., Coronado, R. & Meissner, G. *J. gen. Physiol.* **88**, 573–588 (1986).
7. Ross, C. A. *et al.* *Nature* **339**, 468–470 (1989).
8. Ellisman, M. H. *et al.* *Neuron* **5**, 135–146 (1990).
9. Watras, J., Bezprozvanny, I. & Ehrlich, B. E. *J. Neurosci.* (in the press).
10. Volpe, P., Anderson-Lang, B. H. & Nickols, G. K. *Am. J. Physiol.* **258**, C1077–C1085 (1990).
11. Watras, J. & Benevolensky, D. *Biochim. biophys. Acta* **931**, 354–363 (1987).
12. Ferris, C. D., Hagan, R. L. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2147–2151 (1990).
13. Iino, M. *J. gen. Physiol.* **95**, 1103–1122 (1990).
14. Worley, P. F., Baraban, J. M., Supattapone, S., Wilson, V. S. & Snyder, S. H. *J. biol. Chem.* **262**, 12132–12136 (1987).
15. Varney, M. A., Rivera, J., Lopez Bernal, A. & Watson, S. P. *Biochem. J.* **269**, 211–216 (1990).
16. Danoff, S. K., Supattapone, S. & Snyder, S. H. *Biochem. J.* **254**, 701–705 (1988).
17. Rousseau, E., Smith, J., Henderson, J. & Meissner, G. *Biophys. J.* **50**, 1009–1014 (1986).
18. Borgatta, L., Watras, J., Katz, A. M. & Ehrlich, B. E. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2486–2489 (1991).
19. Airey, J. A. *et al.* *J. biol. Chem.* **265**, 14187–14194 (1990).
20. Supattapone, S., Worley, P. F., Baraban, J. M. & Snyder, S. H. *J. biol. Chem.* **263**, 1530–1534 (1988).
21. Mignery, G. A., Sudhof, T. C., Takei, K. & De Camilli, P. *Nature* **342**, 192–195 (1989).
22. Maeda, N., Niinobe, M. & Mikoshiba, K. *EMBO J.* **9**, 61–67 (1990).
23. Chadwick, C. C., Saito, A. & Fleischer, S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2132–2136 (1990).
24. Mignery, G. A. & Sudhof, T. C. *EMBO J.* **9**, 3893–3898 (1990).
25. Meyer, T., Wensel, T. & Stryer, L. *Biochemistry* **29**, 32–37 (1989).
26. Shoshan-Barmatz, V., Zhang, G. H., Carretson, L. & Kraus-Friedmann, N. *Biochem. J.* **268**, 699–705 (1990).
27. Herrmann-Frank, A., Darling, E. & Meissner, G. *Biophys. J.* **57**, 156a (1990).
28. Berridge, M. J. & Irvine, R. F. *Nature* **341**, 197–205 (1989).
29. Osipchuk, Y. V., Wakui, M., Yule, D. I., Gallacher, D. V. & Peterson, O. H. *EMBO J.* **9**, 697–704 (1990).
30. Meyer, T. & Stryer, L. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5051–5055 (1988).
31. Hartoontumian, A. T., Kao, J. P. Y., Paranjape, S. & Tsien, R. Y. *Science* **251**, 75–78 (1991).
32. Orellana, S., Solski, P. A. & Brown, J. H. *J. biol. Chem.* **262**, 1638–1643 (1987).
33. Hingorani, S., Ondrias, K., Agnew, W. & Ehrlich, B. *Soc. Neurosci. Abstr.* **16**, 1018 (1990).
34. Fabiato, A. & Fabiato, F. *J. Physiol. Paris* **75**, 463–505 (1979).
35. Finch, E. A., Turner, T. J. & Goldin, S. M. *Science* **252**, 443–446 (1991).

ACKNOWLEDGEMENTS. We thank W. K. Chandler, L. Jaffe and L. B. Flash Cohen for comments on the manuscript. This work was supported by the NIH.

Positive feedback in the activation of G1 cyclins in yeast

Léon Dirick & Kim Nasmyth

Research Institute of Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria

YEAST cells become committed to the mitotic cell cycle at a stage during G1 called Start¹. To enter Start, cells must grow to a critical size. They also require the CDC28 protein kinase and at least one of three G1-specific cyclins encoded by *CLN1*, *2*, and *3* (refs 2–4). It is thought that Start is triggered by the accumulation of G1 cyclins that bind to the CDC28 kinase and activate it. So what determines the accumulation of G1 cyclins? For *CLN1* and *CLN2*, transcriptional activation could be involved because their RNAs appear transiently during the cell cycle as cells undergo Start⁵. Here we report that the appearance of *CLN1* and *CLN2* RNAs depends on an active CDC28 kinase and is stimulated by *CLN3* activity. We propose that CDC28 kinase activity due to *CLN1* and *CLN2* proteins arises through a positive feedback loop which allows *CLN* proteins to promote their own synthesis.

Two transcription factors, SWI4 and SWI6, that are required for the Start-dependent activation of the yeast *HO* endonuclease gene^{6,7} are also needed for the activity of all three G1 cyclins⁸.

To explain how SWI4 and SWI6 function can be dependent on the CDC28 kinase (as in the case of *HO*) even though they are activators (through the *CLN* genes) of CDC28, we have proposed that the full activation of CDC28 in G1 involves a positive feedback loop in which SWI4 and SWI6 activate G1 cyclins, which then activate CDC28 kinase which in turn activates SWI4 and SWI6 and so on (see Fig. 1 and its legend). This hypothesis predicts that the accumulation of *CLN1* and *CLN2* transcripts should be CDC28-dependent. We have therefore compared the accumulation of *CLN1* and *CLN2* transcripts in wild-type and temperature-sensitive *cdc28-4* mutants as G1 cells are incubated at the restrictive temperature. Figure 2 shows that the rapid appearance of *CLN1* and *CLN2* RNAs as wild-type cells undergo Start following release from α factor induced G1 arrest does not take place in the *cdc28-4* mutant. Similar results are observed with G1 cells obtained by elutriation (Fig. 3b) or with G1 stationary phase cells inoculated into fresh medium (Fig. 3a). We conclude that the cell cycle-dependent activation of *CLN1* and *CLN2* is indeed dependent on CDC28 activity.

Another prediction of our hypothesis is that the activation of one G1 cyclin should cause the activation of others. The G1 arrest induced by α factor⁹ allows this to be tested. This factor may cause G1 arrest by inhibiting the function of all three G1 cyclin genes¹⁰. The *CLN3-1* mutation, however, causes *CLN3* to become hyperactive and to drive cells through Start even in

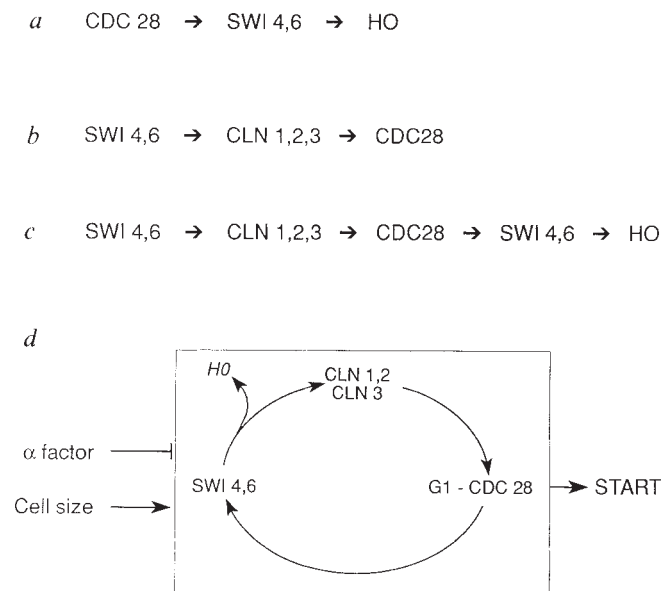
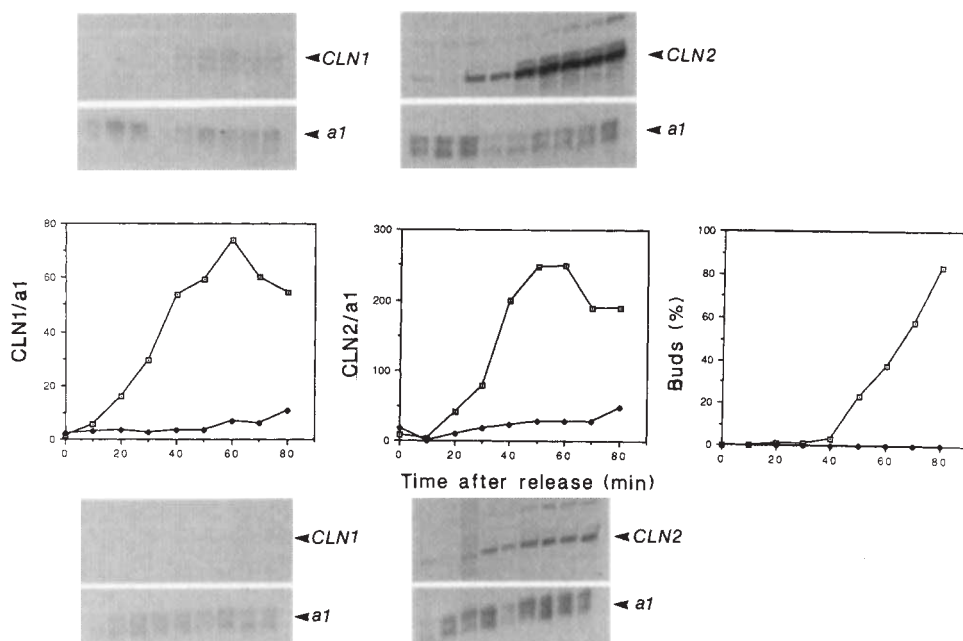


FIG. 1 The positive feedback loop hypothesis. *a*, *HO* transcription depends on the CDC28 protein kinase and only occurs transiently as cells enter Start¹⁶. This behaviour is due to several copies in the *HO* promoter of CACGA₄ sequences to which a complex containing SWI4 and SWI6 proteins bind (refs 7, 17, 18 and M. R. Taba, I. Muroff, D. Lydall, G. Tebb and K.N., manuscript in preparation). *b*, Although neither *SWI4* nor *SWI6* is an essential gene, they share an essential function for cell division because *swi4 swi6* double mutants are inviable. This essential function of SWI4 and SWI6 is to activate the G1 cyclins encoded by *CLN1*, *CLN2*, and *CLN3* (ref. 8), which are thought to be necessary for CDC28 kinase activity in G1. SWI4 and SWI6 appear to be required for *CLN1* and *CLN2* transcription but for the post-translational activity of *CLN3*. *c*, Thus, SWI4 and SWI6 function is not only dependent on the CDC28 kinase (as in the case of *HO*, and outlined in *a*) but also is involved in the activation of the kinase through G1 cyclins, as outlined in *b*. One possible explanation for this conundrum is that SWI4 and SWI6 have two modes of functioning: one CDC28-dependent (downstream of CDC28) and the other CDC28-independent (upstream of CDC28). *d*, Another possibility is that the activation of G1 cyclins takes place through a positive feedback loop. This model leads to two predictions: first, the accumulation of *CLN1* *CLN2* transcripts should be CDC28-dependent (not predicted by the linear model in *c*) and second the activation of one G1 cyclin should lead to the activation of the others. Note that *c* and *d* are not mutually exclusive.

FIG. 2 The cell cycle-dependent appearance of *CLN1* and *CLN2* transcripts depends on CDC28 activity. Exponential cultures of congenic *MATa* wild-type (K699) and *cdc28-4* (K1989) cells growing at 25 °C were arrested in G1 by treatment with 1 $\mu\text{g ml}^{-1}$ α factor for 3 h. Cells were then collected by filtration and resuspended in fresh 37 °C medium lacking α factor. The amount of *CLN1*, *CLN2*, and *MATa1* (as an internal control) transcripts were measured by S1 mapping⁸. The first two graphs show the ratio of *CLN1* (left) and *CLN2* (right) to *MATa1* transcripts ($\times 100$). The bands which were cut out and counted are shown above (wild type) and below (*cdc28-4*) the relevant graphs. The third graph (on the right) shows the percentage of cells possessing a bud. The full genotype of K699 is described in ref. 8. Wild type, $\square$; *cdc28-4*, $\blacklozenge$.

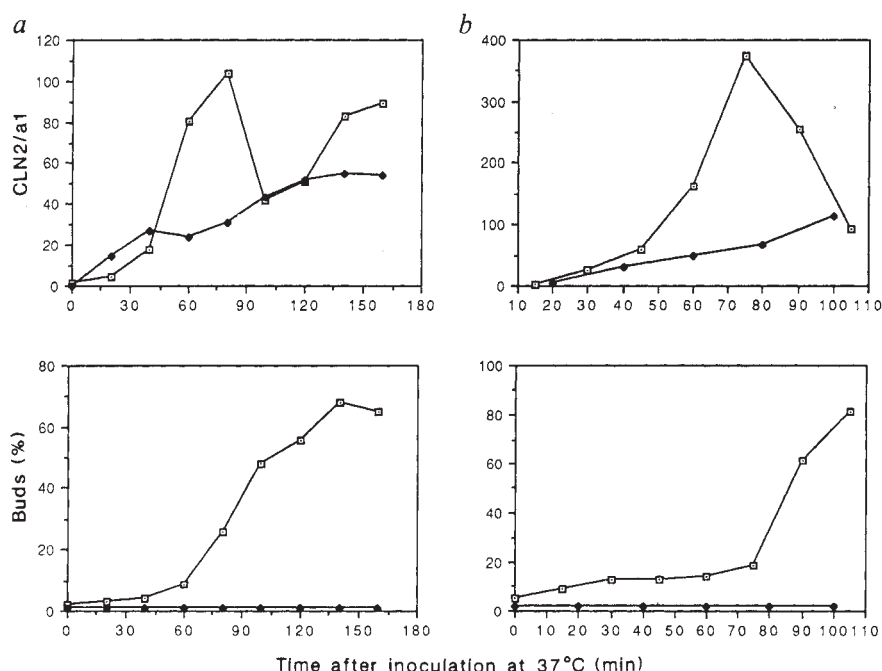


the presence of α factor^{3,11}. To address whether the hyperactivity of *CLN3-1* leads to the activation of *CLN2*, we compared the amount of *CLN2* RNAs after the addition of α factor to asynchronously growing cultures of wild-type and *CLN3-1* mutant cells (Fig. 4A). In wild-type cells, the amount of *CLN2* RNAs drops to less than 10% after 4 h. In *CLN3-1* mutants, a similar drop starts but is followed by the reappearance of *CLN2* RNAs. Thus, the repression of *CLN2* caused by α factor can be overridden by the *CLN3-1* mutation. Activation of *CLN2* transcription by *CLN3-1* actually contributes to the ability of *CLN3-1* mutants to escape from G1 arrest induced by α factor (Fig. 4C). The *cln1 cln2 CLN3-1* mutants are less able to grow on plates containing high concentrations of α factor than *CLN3-1* or even *cln1 CLN3-1* strains, although they are less sensitive than wild-type cells; a result that is confirmed by halo assays of cells' α -factor sensitivity.

We conclude that the passage through Start of *CLN3-1* mutants in the presence of high levels of α factor depends on the activation of *CLN2*. Thus, the activity of one G1 cyclin (in this case *CLN3-1*) leads to the activity of another (*CLN2*) and only thereby to the passage of cells through Start.

We also analysed *CLN2* transcripts when *CLN1* and *CLN2* contain mutations and the cells' viability is due to expression of *CLN3* from the *GAL1-10* promoter¹¹ (Fig. 4B). Addition of glucose to such cells represses *CLN3* transcription and leads to uniform G1 arrest after 100 min. The amount of *CLN2* transcripts rapidly drops sixfold. By contrast, the amount of *CLN2* RNAs remains constant in wild-type cells following such a shift. The amount of *CLN2* RNAs in cells expressing *CLN3* from the *GAL* promoter is in fact much higher than in wild-type cells when grown in galactose. We do not know whether this is due to hyperactivation of *CLN2* transcription due to higher than

FIG. 3 CDC28-dependent and -independent appearance of *CLN2* transcripts following the inoculation of stationary phase G1 cells into fresh medium and from G1 cells prepared by elutriation. *a*, Wild-type (K699) and *cdc28-4* (K1989) cells were grown to stationary phase on YEPD plates at 25 °C as previously described¹⁷ and then inoculated into fresh YEPD medium at 37 °C. The amount of *CLN2* and *MATa1* (as an internal control) transcripts were measured by quantitative S1 mapping. The *MATa1* transcript varies little during the time course and the top graph shows the ratio of *CLN2* to *MATa1* RNAs ($\times 100$). The percentage of cells with buds is shown underneath. Wild type, $\square$; *cdc28-4*, $\blacklozenge$. *b*, G1 populations prepared by centrifugal elutriation (M. R. Taba, I. Muroff, D. Lydall, G. Tebb and K.N., manuscript in preparation) from wild-type (K699) and *cdc28-4* (K1989) cells growing in YEPD at 25 °C were incubated at 37 °C. *CLN2* and *MATa1* RNAs were measured by quantitative S1 mapping. The *MATa1* transcript remains constant and is used as an internal control. The graphs show the ratio of *CLN2* to *MATa1* RNAs $\times 100$ (top) and the percentage of cells with buds (bottom). Due to the difficulty of isolating sufficient quantities of G1 cells, about 5% of the starting culture of wild-type cells were already in S phase and had buds. There was no such contamination of the *cdc28-4* cells.



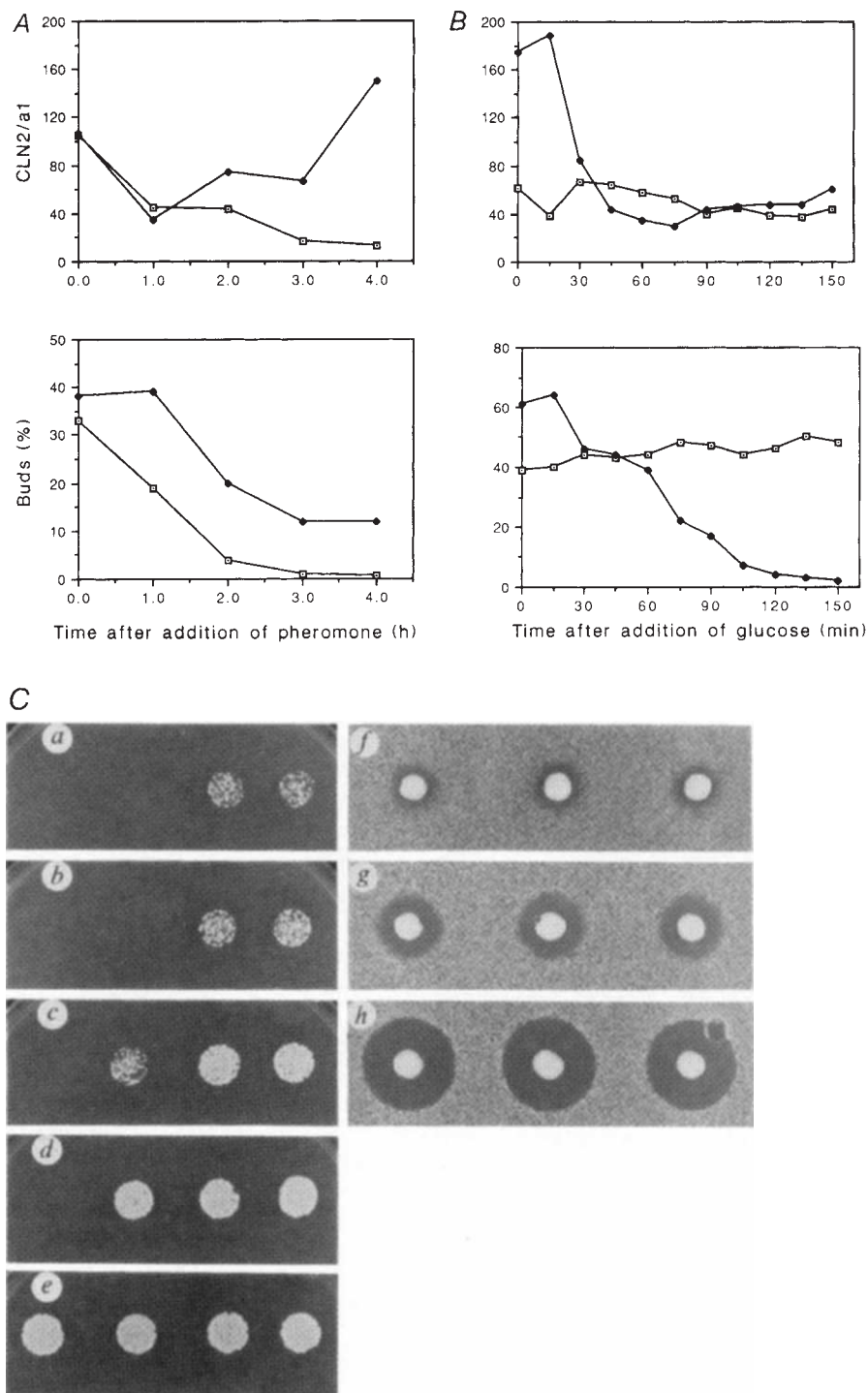
normal levels of CLN3 protein or whether it is due to the *CLN2* RNAs being truncated and having a higher stability (see legend to Fig. 4A). In either case, it is clear that under these circumstances the abundance of *CLN2* transcripts is dependent on the activity of CLN3.

The *cdc28-4* mutation does not totally abolish the accumulation of *CLN2* transcripts at the restrictive temperature. The rapid rise in *CLN2* transcripts in wild-type G1 cells obtained by elutriation or starvation is replaced by a slower and more linear rise in *cdc28-4* mutants (Fig. 3). This may indicate that the *cdc28-4* allele is leaky and that the mutant CDC28 protein can still partially stimulate the amount of *CLN2* RNA. Another explanation is that there exists a second pathway for *CLN2* activation that is *CDC28*-independent. Leakiness is inconsistent

with the observation that none of the mutant cells bud or express *HO* RNA (Fig. 3 and refs 6, 7).

The proposal that a positive feedback loop is involved in the activation of the *CDC28* protein kinase in G1 may help to explain the irreversibility of Start. A property of positive feedback loops is that they have only two stable states: on and off. We propose that the full activation of the *CDC28* kinase due to positive feedback could be one of the molecular bases behind the commitment of cells to the mitotic cell cycle at Start (see also ref. 12). Our proposal that there is also a *CDC28*-independent pathway of *CLN2* activation may explain why *CLN1* and *CLN2* RNAs only appear after cells reach a certain size^{5,13}. The *CDC28*-independent activation pathway may steadily increase in strength as G1 cells grow, eventually becoming strong enough

FIG. 4 The interdependence of CLN activities. A, The effect of α factor on the amount of *CLN2* transcripts in wild-type and *CLN3-1* mutants with an α mating type. α factor (final concentration $1 \mu\text{g ml}^{-1}$) was added to exponential cultures of congeneric *bar1* wild-type (K2149) and *CLN3-1* mutant (K2184) cells and *CLN2* and *MATa1* transcripts were measured by quantitative S1 mapping at 60 min intervals after α -factor addition. The top graph plots the ratio of *CLN2* to *MATa1* transcripts ($\times 100$) in wild type and *CLN3-1*. The percentage of budded cells in each culture is shown in the graph underneath. Wild-type cells remain arrested in G1 until they die. *CLN3-1* mutants partially arrest and then resume division (data not shown). Both K2149 and K2184 are congenic with K1107 whose full genotype is described in ref. 8. Wild type, $\square$; *CLN3-1*, $\blacklozenge$. B, The effect on amounts of *CLN2* transcript of repressing *GAL1-10* driven *CLN3* transcription from a plasmid in a strain (K2384) whose endogenous *CLN1*, *CLN2*, and *CLN3* genes have been inactivated by deletion. Wild-type (K1107) and *GAL-CLN3* cells (K2384) growing at 30 °C in YEP medium containing 2% raffinose and 2% galactose were collected by filtration and resuspended in fresh YEP medium containing 2% raffinose and 2% glucose. The top graph shows plots the ratio of *CLN2* to *MATa1* transcripts as measured by S1 mapping. Strain K2384 contains a deletion of *CLN2* which removes DNA between the *XhoI* and *SacI* sites in the gene. As the probe only contains sequences 5' to the *XhoI* site, the DNA protected from S1 nuclease is identical to that seen in wild type. The *CLN2* transcript made in the cell will contain the wild-type 5' and 3' sequences but lacks a large fraction of the internal sequences, which could affect its stability. K2384 is congenic with K1107 and contains the *cln1::hisG*, *cln2::del* and *cln3::LEU2* deletions described in ref. 8. *CLN3* is expressed from the *GAL1-10* promoter on a YCp50 derived plasmid¹⁹. C, The α factor resistance of *CLN3-1* depends partially on *CLN2*. a–e, α factor sensitivity of *MATa bar1* cells monitored by adding 250 cells in a 5 μl drop to YEPD plates containing 600 (a), 150 (b), 37.5 (c), 9.38 (d) and 2.34 nM (e) α factor. The four different strains were from left to right: wild type (K2149), *cln1 Δ cln2 Δ CLN3-1* (K2301), *cln1 Δ CLN3-1* (K2282), and *CLN3-1* (K2184). The plates were incubated for two days at 30 °C. f–h, α -factor sensitivity of *MATa bar1* cells monitored by halo assays. *MATa* cells (K217) were spotted in triplicate on soft agar layers containing 2.5×10^5 *MATa bar1* cells of the following strains: *cln1 Δ CLN3-1* (K2282) (f), *cln1 Δ cln2 Δ CLN3-1* (K2301) (g) and wild type (K2149) (h). The halo of *CLN3-1* cells is identical to that of *cln1 CLN3-1* cells (data not shown). All strains are congenic with K1107 and the *cln* deletions are described in ref. 8.



to trigger the full activation of *CLN1* and *CLN2* through the positive feedback loop.

The physiological significance of the loop is still unclear because *cln1 cln2 cln3* triple mutants kept alive by virtue of *CLN2* expressed from a constitutively expressed promoter can proliferate efficiently^{2,8}. This means either that the regulation of G1 cyclins is not essential for viability or that there exists also a post-transcriptional mechanism for *CLN2*'s regulation, as there must be for *CLN3* (refs 3, 11). It is also unknown how the yeast pheromone signal transduction pathway prevents the loop from firing and how the loop is eventually turned off to make way for the formation of different forms of the CDC28 kinase (for example involving G2-specific cyclins)^{14,15}.

A question remains concerning the mechanism by which CDC28 kinase activity raises *CLN1* and *CLN2* transcript levels. SWI4 and SWI6, which are components of a transcription complex required for CDC28-dependent *HO* expression, are also required for *CLN2* transcription. Therefore, one possibility is that phosphorylation of SWI4 and SWI6 complexes by the CDC28 protein kinase may increase their ability to activate transcription. □

Received 9 May; accepted 3 June 1991.

1. Pringle, J. R. & Hartwell, L. H. In *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*. 97–142 (Cold Spring Harbor Laboratory, New York, 1981).
2. Richardson, H. E., Wittenberg, C., Cross, F. & Reed, S. I. *Cell* **59**, 1127–1133 (1989).
3. Nash, R., Tokowa, G., Anand, S., Erickson, K. & Futcher, A. B. *EMBO J.* **7**, 4335–4346 (1988).
4. Cross, F. R. *Molec. cell. Biol.* **10**, 6482–6490 (1990).
5. Wittenberg, C., Sugimoto, K. & Reed, S. I. *Cell* **62**, 225–237 (1990).
6. Stern, M., Jensen, R. & Herskowitz, I. *J. molec. Biol.* **178**, 853–868 (1984).
7. Breeden, L. & Nasmyth, K. *Cell* **48**, 389–397 (1987).
8. Nasmyth, K. A. & Dirick, L. *Cell* (in the press).
9. Cross, F., Hartwell, L. H., Jackson, C. & Konopka, J. B. *A. Rev. Cell Biol.* **4**, 429–457 (1988).
10. Nasmyth, K. A. *Cell* **63**, 1117–1120 (1990).
11. Cross, F. R. *Molec. cell. Biol.* **8**, 4675–4684 (1988).
12. Reed, S. I. *Trends Genet.* **7**, 95–99 (1991).
13. Price, C., Nasmyth, K. & Schuster, T. *J. molec. Biol.* **218**, 543–556 (1991).
14. Surana, U. *et al.* *Cell* **65**, 145–162 (1991).
15. Ghiara, J. B. *et al.* *Cell* **65**, 163–174 (1991).
16. Nasmyth, K. A. *Nature* **302**, 670–676 (1983).
17. Nasmyth, K. *Cell* **42**, 225–235 (1985).
18. Andrews, B. J. & Herskowitz, I. *Nature* **342**, 803–833 (1989).
19. Nasmyth, K. A., Adolf, G., Lydall, D. & Seddon, A. *Cell* **62**, 631–647 (1990).

Acquired immunity and epidemiology of *Schistosoma haematobium*

M. E. J. Woolhouse*, P. Taylor†, D. Matanhire‡ & S. K. Chandiwan‡

* Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

† Training Centre for Water and Sanitation, University of Zimbabwe, PO Box MP167, Harare, Zimbabwe

‡ Blair Research Laboratory, PO Box 8105, Causeway, Harare, Zimbabwe

HUMAN immune responses to schistosome infection have been characterized in detail^{1–5}. But there has been controversy⁶ over the relative importance of ecological factors (variation in exposure to infection) and immunological factors (acquired immunity) in determining the relationships between levels of infection and age typically found in areas where infection is endemic^{7,8}. Independent effects of exposure and age on the rates of reinfection with *Schistosoma haematobium* after chemotherapy have been demonstrated in the Gambia⁹ and Zimbabwe⁸. This age effect could be the result of acquired immunity to infection. Indeed, allowing for variation in exposure and age, low rates of reinfection in the Gambia are correlated with high amounts of specific IgE antibodies¹⁰—human IgE can kill *S. mansoni* schistosomulae *in vitro*¹¹. Further, animals can acquire immunologically mediated resistance to *S. mansoni* infection^{12–14}, although nonimmunological factors could also be

involved¹⁵. Acquisition of this immunity seems to be related to the cumulative effects of repeated infection and provides only partial protection^{12,13,16}. These characteristics are consistent with immuno-epidemiological data for both *S. mansoni* and *S. haematobium* infections of humans^{4,17}. We have now analysed age-prevalence data for human infection with *S. haematobium*, and find patterns of variation that are indeed consistent with the epidemiological effects of acquired immunity predicted by mathematical models.

The epidemiological effects of an immune response to schistosome infection can be investigated using a mathematical model developed by Anderson and colleagues^{18,19}. This is an age-structured immigration-death model that considers changes with age, a , in the mean intensity of infection (number of adult schistosomes), $m(a)$, as a function of the per capita rate of infection, Λ , and the per capita parasite mortality rate, γ . Acquired immunity moderates the rate of parasite establishment. The model predicts the following consequences of a higher rate of infection, Λ , on endemic epidemiological patterns: (1) an increase in the convexity of the age-intensity curve; (2) an increase in the peak intensity of infection, $m(a^*)$; and (3) for certain parameter values, a decrease in the age of peak intensity of infection, a^* . The last two of these imply a negative correlation between $m(a^*)$ and a^* for populations with different values of Λ (Fig. 1). These predictions hold even if Λ is itself age-dependent, perhaps owing to age-related variation in exposure, provided that the form of the relationship does not vary systematically between populations subject to high or low

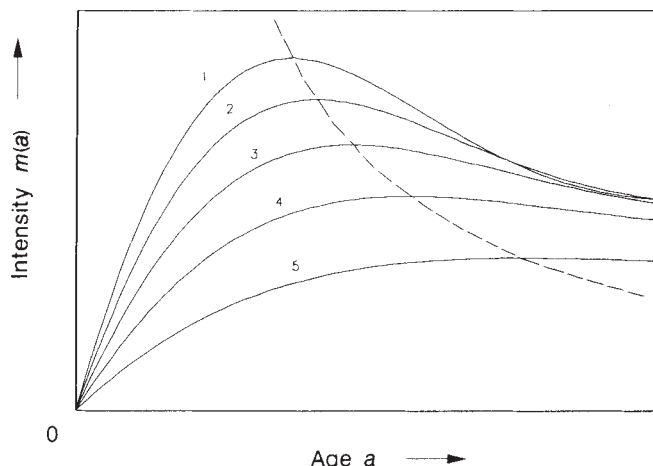


FIG. 1 Predicted patterns for age-related changes in the intensity of infection using a version of the model of Anderson and colleagues^{18,19}. The model considers changes in the mean intensity of infection, m , with respect to age, a . For certain parameter values (see below) the model predicts an age-intensity curve of the form:

$$m(a) = A + e^{-(\gamma + \sigma)a/2} [2\Lambda \sin(a\theta/2) [1 - A(\gamma + \sigma)/2\Lambda] / \theta - A \cos(a\theta/2)] \quad (1)$$

with $A = \Lambda\sigma/(\sigma\gamma + \nu\Lambda)$ and $\theta = [4\nu\Lambda - (\gamma - \sigma)^2]^{1/2}$, where Λ is the per capita rate of infection, ν is a measure of the decrease in the per capita rate of parasite establishment due to acquired immunity, σ is the rate of loss of immunological memory, and γ is the per capita mortality rate for adult parasites. The parameter Λ is a function of the mean per capita fecundity of adult parasites, intermediate host snail population density, life expectancy of infected snails, human population density, per capita rate of infection of snails per adult parasite and per capita rate of infection of humans per infected snail. The rates of infection of both humans and snails are functions of the mean water contact rate of the human population. Equation (1) applies where parameter values satisfy the condition $4\nu\Lambda > (\gamma - \sigma)^2$. The solid curves in the figure represent equilibrium age-intensity curves predicted by equation (1) with parameter values $\nu = 0.0015$, $\sigma = 0.04$, $\gamma = 0.027$ throughout and $\Lambda = 2.2, 1.8, 1.4, 1.0$ and 0.6 for curves labelled 1–5, respectively. The shift in the predicted peak intensity of infection, $m(a^*)$ with the predicted age of peak intensity, a^* , is shown by the broken curve. This relationship is generated solely by variation in Λ .

transmission, that is, if $d\Lambda/da$ is independent of Λ (see below). The model formalizes earlier predictions based on verbal arguments²⁰.

Variations in age-intensity patterns for human infection with schistosomes^{20,21} and hookworm¹⁹ have been reported to be consistent with these predictions, but a robust analysis has not been possible owing to the lack of suitable field data. We have analysed 17 data sets recording the prevalence within age groups, $P(a)$, of *S. haematobium* infection for more than 2,000 children in Zimbabwe. There is a statistically significant negative correlation between $P(a^*)$ and a^* across these data sets (Fig. 2). The relationship between prevalence, P , and m depends on the distribution of schistosomes among hosts (the more aggregated this distribution, the lower P is for a given m), so the pattern shown in Fig. 2 could reflect differences in the parasite distribution rather than in mean intensities of infection. But available data indicate that the degree of aggregation, as indexed by the variance-to-mean ratio for egg counts, does not increase with age (Fig. 3), so the pattern of Fig. 2 must reflect differences in $m(a^*)$.

The observed pattern is therefore consistent with model predictions. The only other interpretation possible is that there is a relationship between Λ and age, perhaps due to age-related variation in exposure, such that populations with high overall Λ also tend to have peak values of Λ at an earlier age. None of the many epidemiological and sociological studies carried out in Zimbabwe indicates such a relationship^{22,23}.

Although the analysis captures the essentials of the dynamic interaction between infection and acquired immunity, it ignores certain features. We have not considered the effects of blocking antibodies, which are thought to be involved in human schistosome infection, or age-dependency in the immune response

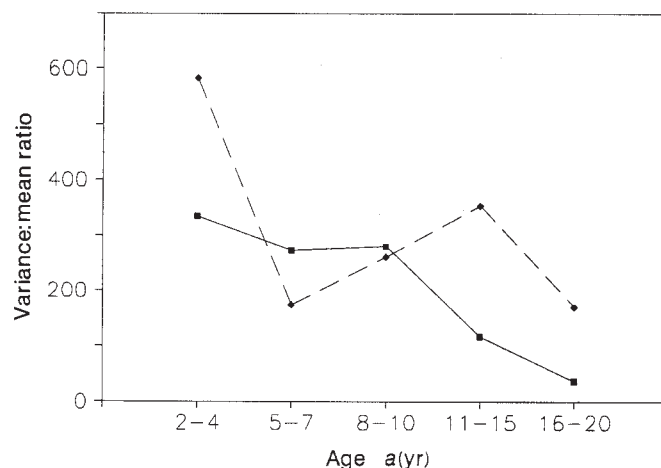


FIG. 3 Relationships between the variance-to-mean ratio for intensity of *S. haematobium* infections and age. The data are egg counts per 10 ml urine (single sample) from children between 6 and 20 years old, divided into five age classes. ■, Sample of 237 children from a community in eastern Zimbabwe⁸. ◆, Sample of 188 children from a community in the Zimbabwe highveld (S.K.C. and M.E.J.W., unpublished data). Kendall's rank correlation $\tau = -0.800$ ($P = 0.04$) and -0.400 ($P = 0.24$), respectively. Peak levels of infection occur among 8–10-year-old children in both samples.

itself²⁴. More information is needed before these effects can be incorporated in mathematical models. Nor have we considered the effects of acquired immunity on schistosome mortality and fecundity and on the fraction of eggs passed out of the host's body, effects which have been reported in animal models^{13,25}. Because egg output is the most widely used index of the intensity of infection in humans, any effects of acquired immunity on egg output need to be better understood if immuno-epidemiological data are to be interpreted correctly.

In conclusion, we have shown patterns of variation in age-prevalence data for human infection with *S. haematobium* that are consistent with current understanding of the epidemiological effects of acquired immunity. To our knowledge, such a relationship has not previously been demonstrated for any human macroparasite infection. The interaction between schistosome infection and acquired immune responses are complex and we expect that mathematical models will prove valuable for the interpretation of field data. □

Received 22 March; accepted 24 April 1991.

1. Sturrock, R. F. et al. *Trans. R. Soc. trop. Med. Hyg.* **77**, 363–371 (1983).
2. Butterworth, A. E. et al. *Trans. R. Soc. trop. Med. Hyg.* **79**, 393–408 (1985).
3. Butterworth, A. E. et al. *Parasitology* **94**, 281–300 (1987).
4. Hagan, P. et al. *Trans. R. Soc. trop. Med. Hyg.* **81**, 938–946 (1987).
5. Dunne, D. W. et al. *Eur. J. Immun.* **18**, 123–131 (1988).
6. Wilkins, H. A. in *Biology of Schistosomes* (eds Rollinson, D. & Simpson, A. J. G.) 379–398 (Academic, London, 1987).
7. Wilkins, H. A., Goll, P. H., Marshall, T. F. & Moore, P. J. *Trans. R. Soc. trop. Med. Hyg.* **78**, 216–221 (1984).
8. Chandiwana, S. K., Woolhouse, M. E. J. & Bradley, M. *Parasitology* **102**, 73–83 (1991).
9. Wilkins, H. A., Blumenthal, U. J., Hagan, P., Hayes, R. J. & Tulloch, S. *Trans. R. Soc. trop. Med. Hyg.* **81**, 29–35 (1987).
10. Hagan, P., Blumenthal, U. J., Dunn, D., Simpson, A. J. G. & Wilkins, H. A. *Nature* **349**, 243–245 (1991).
11. Capron, A., Dessaint, J. P., Capron, M., Ouma, J. H. & Butterworth, A. E. *Science* **238**, 1065–1072 (1987).
12. Smithers, S. R. & Terry, R. J. *Trans. R. Soc. trop. Med. Hyg.* **61**, 517–533 (1967).
13. Dean, D. A. *Exp. Parasit.* **55**, 1–104 (1983).
14. Pearce, E. J. & McLaren, D. J. *Parasitology* **87**, 465–479 (1983).
15. Wilson, R. A. *Parasit. Today* **6**, 354–358 (1990).
16. James, S. L. & Cheever, A. W. *Parasitology* **91**, 301–315 (1985).
17. Hagan, P. in *Biology of Schistosomes* (eds Rollinson, D. & Simpson, A. J. G.) 295–320 (Academic, London, 1987).
18. Crombie, J. A. & Anderson, R. M. *Nature* **315**, 491–493 (1985).
19. Anderson, R. M. & May, R. M. *Nature* **315**, 493–496 (1985).
20. Clarke, V. de V. C. *Afr. J. Med.* **12**, Suppl., 1–30 (1966).
21. Anderson, R. M. in *Clinical Tropical Medicine and Communicable Diseases* Vol. 2 (ed. Mahmoud, A. A. F.) 279–300 (Bailliere Tindall, London, 1987).
22. Chandiwana, S. K. *Soc. Sci. Med.* **25**, 495–505 (1987).
23. Chandiwana, S. K. & Christensen, N. O. *Trop. med. Parasit.* **39**, 187–193 (1988).
24. Butterworth, A. E. & Hagan, P. *Parasit. Today* **3**, 11–16 (1987).

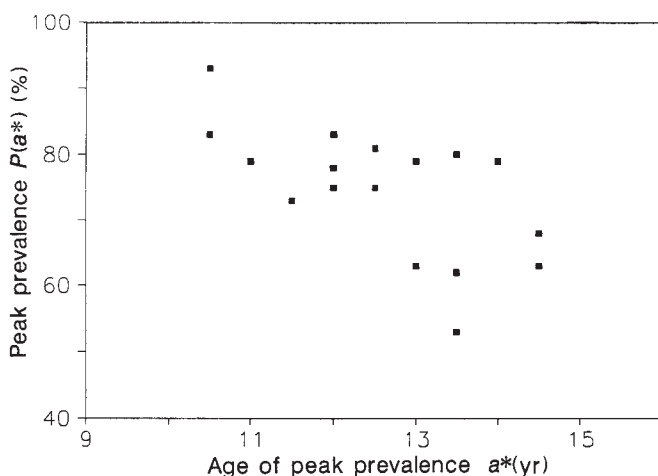


FIG. 2 Relationship between peak prevalence of infection, $P(a^*)$, and age of peak prevalence, a^* , for *S. haematobium* infections of school-children in Zimbabwe. Data are from 17 primary schools in the Zimbabwe highveld. A total of 2,185 children between 6 and 16 years old were screened for infection by examination of a 10-ml urine sample^{26,27}. Data from adjacent one-year age classes have been pooled where necessary to provide estimates of $P(a^*)$ based on more than 20 individuals. In no case where data were available from additional age classes (six schools) was peak prevalence observed outside the age range 6–16 years. The schools were located within an area $\sim 100 \times 100$ km and represent contiguous populations within a geographically, culturally and sociologically homogeneous region^{26,27} (compare with refs 19 and 21). Several factors may contribute to variation in the per capita rates of infection experienced by these different populations: differences in human population densities, differences in population densities of the intermediate host snails (*Bulinus globosus*), and differences in mean water contact rates (reflecting differences in the availability and convenience of natural water bodies). The relationship between $P(a^*)$ and a^* is in the direction predicted by the model incorporating acquired immunity (Fig. 1) and is statistically significant: Kendall's rank correlation $\tau = -0.438$, $n = 17$, $P < 0.01$ (one-tailed test).

25. Doenhoff, M. J., Hassounah, O., Murare, H., Bain, J. & Lucas, S. *Trans. R. Soc. trop. Med. Hyg.* **80**, 503–514 (1986).
26. Taylor, P., Chandiwana, S. K. & Matanhire, D. *Acta Trop.* **47**, 91–100 (1990).
27. Taylor, P., Chandiwana, S. K. & Matanhire, D. *Community Based Schistosomiasis Control in Zimbabwe* (IDRC, Canada, in the press).

ACKNOWLEDGEMENTS. We thank R. M. Anderson, A. E. Butterworth, R. D. Gregory, P. Hagan, A. E. Keymer, A. R. McLean & H. A. Wilkins for discussions. This work was supported by the International Development Research Centre, Canada, and a Beit Memorial Fellowship for Medical Research (M.E.J.W.).

Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation

Jean-Paul Issartel, Vassilis Koronakis & Colin Hughes

Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

HAEMOLYSIN secreted by pathogenic *Escherichia coli* binds to mammalian cell membranes, disrupting cellular activities and lysing cells by pore-formation. It is synthesized as nontoxic prohaemolysin (proHlyA), which is activated intracellularly by a mechanism dependent on the cosynthesized HlyC. Haemolysin is one of a family of membrane-targeted toxins, including the leukotoxins of *Pasteurella* and *Actinobacillus* and the bifunctional adenylate cyclase haemolysin of *Bordetella pertussis*, which require this protoxin activation^{1–5}. HlyC alone cannot activate proHlyA, but requires a cytosolic activating factor⁶. Here we report the cytosolic activating factor is identical to the acyl carrier protein and that activation to mature toxin is achieved by the transfer of a fatty acyl group from acyl carrier protein to proHlyA. Only acyl carrier protein, not acyl-CoA, can promote HlyC-directed proHlyA acylation, but a range of acyl groups are effective.

In vitro post-translational activation of proHlyA to the haemolytically active toxin is achieved by mixing cell extracts containing separately the proHlyA of relative molecular mass 110,000 (M_r 110K) and the 19.8K HlyC, but activation is lost when both proteins are purified⁶. Activation is recovered by addition of an *E. coli* cell extract, demonstrating the existence of an essential cytosolic activating factor (CAF)⁶. Purification of the protease-sensitive CAF increased about 120-fold in the specific proHlyA-activating activity (Fig. 1a, b). Anion exchange chromatography revealed that CAF is highly negatively charged and gel filtration indicated that it has an M_r of 9–10K (Fig. 1a). SDS-PAGE showed a single polypeptide (Fig. 1b, lanes 1–3) migrating as a diffuse, weakly stained band with a higher apparent M_r of 16–19K. N-terminal sequencing of purified CAF showed that the first 20 amino acids (Fig. 1c) are identical to those of *E. coli* acyl carrier protein (ACP), an 8.86K negatively charged polypeptide which, like CAF, migrates aberrantly during SDS-PAGE^{7,8} (Fig. 1b, lane 4).

In *E. coli*, ACP can function either in the unacylated form, as in the synthesis of periplasmic glucans⁹, or in the acylated form, donating acyl residues linked by a thioester bond to the covalently bound phosphopantotheine group, for example in lipid biosynthesis. That proHlyA activation requires acyl-ACP was indicated by the loss of CAF activity after treatment with hydroxylamine, dithiothreitol (DTT) or NaBH_4 (not shown), each of which hydrolyses thioester linkages¹⁰ between acyl groups and holo-ACP (ACP-SH). Further evidence was based on the characteristic behaviour of ACP-SH and acyl-ACP during chromatography in octyl sepharose. Although ACP-SH does not bind to this medium the acyl-ACP adsorbs tightly until eluted by 2-propanol¹¹. Purified CAF adsorbed to octyl sepharose, and when eluted by 2-propanol it was able to activate proHlyA in the presence of HlyC (Table 1, bound fraction).

TABLE 1 Binding of CAF to octyl sepharose

	Preacylation (HU)	Postacylation (HU)
CAF	1.2×10^3	1.1×10^5
Unbound fraction	2.0	2.7×10^4
Bound fraction	0.6×10^3	0.9×10^4

A 100 μl aliquot of CAF (purified as in Fig. 1) able to generate 1.2×10^3 HUs was mixed with 20 μl octyl sepharose CL4B (Pharmacia). After incubation for 10 min at 22 °C, the chromatographic medium was pelleted by centrifugation and the supernatant withdrawn (the 'unbound fraction'), the pellet was then mixed with 100 μl HED buffer containing 25% 2-propanol. The supernatant obtained after centrifugation of this was the 'bound fraction'. Both fractions were assayed for proHlyA activation before and after enzymatic acylation (pre- and postacylation, respectively), and the activities expressed in HUs, standardized to 100 μl of the original fractions. Enzymatic acylation was performed essentially as recommended¹²; aliquots of purified CAF, unbound fraction and bound fraction were lyophilized and incubated with 160 μM myristic acid (sodium salt) and 0.8 mU acyl-ACP synthetase (Sigma) in 20 μl . After 30 min at 37 °C the proHlyA activation competence of the samples was assayed. *In vitro* activation of proHlyA in the presence of HlyC and the CAF-containing samples was performed as described⁶. Routinely, purified proHlyA (6 μg) was preincubated for 10 min at 22 °C in the presence of 80 ng purified HlyC and different amounts of CAF in a total of 50 μl HED buffer. Haemolytic activities were assayed by diluting the reaction mixes with 1 ml saline buffer, 150 mM NaCl, 20 mM CaCl_2 containing 2% (v/v) equine erythrocytes. After 30 min at 42 °C debris was cleared by centrifugation, and the haemoglobin released by erythrocyte lysis was read at 543 nm. One HU is defined as the amount of active HlyA promoting the release of 1 nmol haemoglobin per h.

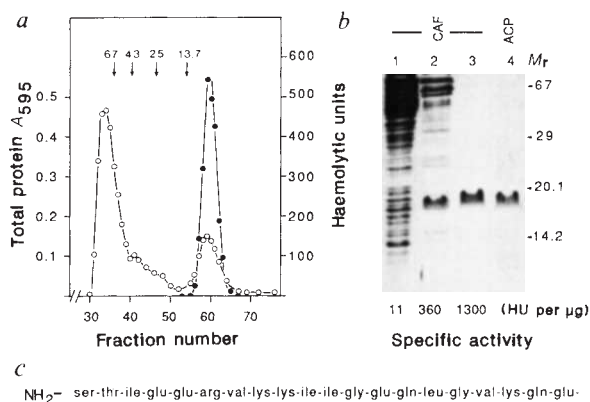


FIG. 1 Identity of the cytosolic activating factor (CAF) with acyl carrier protein (ACP). a, Final purification step of CAF through a G75 Sephadex column; protein elution profile (○), proHlyA activation in haemolytic units (●). Arrows correspond to the elution volumes of M_r markers ($M_r \times 10^{-3}$). b, Analysis by 20% SDS-PAGE. Lanes 1, cytosolic extract 30 μg ; 2, active fractions from Q sepharose chromatography 15 μg ; 3, active fractions from G75 Sephadex 9 μg ; 4, ACP (Sigma) 6 μg . Specific CAF activities are given as HU per μg protein. c, The N-terminal amino-acid sequence of purified CAF obtained by an Applied Biosystems 477A Protein Sequencer after electroblotting onto Problott membrane.

METHODS. *E. coli* strain PR7 was grown in aerated 2TY medium at 37 °C up to an absorbance at 600 nm of 1.6. Pelleted cells (10,000g, 10 min) were resuspended in 1/20 original culture volume then disrupted in a French Press (12,000 p.s.i.). Lysate was centrifuged (30,000g, 15 min) and 10 ml supernatant were loaded on a Q sepharose fast-flow anion exchange chromatography column (Pharmacia, 10 ml), equilibrated in 25 mM HEPES buffer, pH 8.0, 5 mM EDTA, 1 mM DTT (HED buffer). The column was washed with 50 ml HED containing 50 mM NaCl and elution was achieved using a linear 50–500 mM NaCl gradient in 250 ml HED. CAF was assayed by *in vitro* HlyC-dependent activation of proHlyA, as described in the legend to Table 1. Active fractions eluted close to 300 mM NaCl were lyophilized, solubilized with 3 ml HED containing 3 mM magnesium acetate and 70 mM NaCl and loaded on a G75 Sephadex superfine column (Pharmacia, 1.5 × 90 cm) equilibrated in the same buffer. Flow rate was 8 ml h⁻¹ and 2 ml fractions were collected. Proteins were assayed either with a Coomassie-blue staining solution (a) or the more accurate microBCA assay (Pierce) (b).

TABLE 2 Activity of specific acylated forms of ACP in the activation *in vitro* of proHlyA

Fatty acid	None	C ₁₂	C ₁₄	C ₁₆	C _{16:1}	C _{18:1}
Specific activity (HU μg^{-1})	6	1.5×10^4	7.6×10^4	2.7×10^4	3.8×10^4	0.7×10^4
Relative activity	—	0.20	1.0	0.36	0.50	0.09

HlyC-dependent proHlyA activation by ACP (Sigma) enzymatically acylated by various fatty acids, as described in Table 1. ACP (15 μM) was incubated with 160 μM fatty acid (sodium salt) and 1.5 mU acyl-ACP synthetase (Sigma) in 40 μl , at 37 °C for 30 min. Reaction mixes were diluted 500-fold and aliquots were assayed for proHlyA activation as described in Table 1. Haemolytic activities were corrected to take into account that the yields of acylated ACPs are dependent on the nature of the fatty acid used as a substrate. Accordingly, HUs were multiplied by the following correcting factors drawn from the published data¹¹: C₁₂ (lauric acid)=2.2; C₁₄ (myristic acid)=1.1; C₁₆ (palmitic acid)=1.0; C_{16:1} (*cis*-9-hexadecenoic acid/palmitoleic acid)=1.8; C_{18:1} (*cis*-11-octadecenoic acid/*cis*-vaccenic acid)=2.0. Corrected specific haemolytic activities given in the table were used to define the relative ability of the different acylated-ACP to activate proHlyA, taking C₁₄-ACP as a reference.

The unbound fraction was unable to activate proHlyA, but could be converted into the activating form by *E. coli* acyl-ACP synthetase, which catalyses the ligation of fatty acids to ACP-SH in the presence of ATP-Mg²⁺. This enzymatic acylation resulted in about a 10⁴-fold increase in proHlyA activation. CAF and the octyl sepharose-bound fraction were also enhanced, but by 92-fold and 15-fold, respectively. The data suggest that purified CAF contains both acyl-ACP and ACP-SH, and that proHlyA activation requires acyl-ACP. The requirement is strict; neither acyl-CoAs nor free fatty-acids promoted activation. Commercial ACP (ACP-SH) supported low amounts of HlyC-dependent proHlyA activation (specific activity 6 haemolytic units (HU) μg^{-1} ; compare with CAF 1,300 HU μg^{-1}). After enzymatic acylation with a range of fatty acids (C₁₂ to C_{18:1}), the resulting acylated ACPs all had a specific activity about 10⁴-fold higher than ACP-SH (Table 2), with myristoylated ACP the most active. Myristic acid is not as abundant in *E. coli* as palmitic, palmitoleic and *cis*-vaccenic acids, so myristoyl-HlyA may be the most active acylated toxin among a mixture secreted by the cell.

We demonstrated directly that activation involves transfer of the acyl group from acyl-ACP to proHlyA by using enzymatically synthesized [¹⁴C]palmitoyl-ACP. *In vitro* activation to the haemolytic HlyA occurred concomitantly with the radiolabelling of proHlyA, that is by covalent acylation with the [¹⁴C]palmitoyl group (Fig. 2b, lane 2). No stable intermediate fatty-acyl-HlyC was found. No activation and no radiolabelling of proHlyA occurred in the absence of HlyC (Fig. 2b, lane 1).

Acylation is involved in the maturation of many proteins of both prokaryotic and eukaryotic cells, including viral oncogene products, but it is achieved by various mechanisms differing according to the fatty acid transferred, the amino acid modified, and the fatty acyl donor¹². Myristic and palmitic acids are the most common fatty acids cross-linked to proteins. Myristic acid is attached through an amide linkage to an N-terminal glycine by the specific enzyme *N*-myristoyl transferase^{10,13}, but acylation can also occur on internal cysteine or threonine residues (S- and O-ester linkages), and probably also lysines (amide linkage) in processes which do not show strong selectivity for specific fatty acids¹². Fatty acyl-CoAs act as donors in both myristoylation and ester-type acylation. A further hydrophobic modification of eukaryotic proteins is the covalent attachment of glycolipid to the C-terminus of protein (glypiation)¹⁴, whereas lipoproteins sorted to the bacterial outer membrane (for example penicillinase and pullulanase) undergo complex processing in which an *N*-acyl diglyceride cysteine is generated at the N-terminus, using fatty acids derived from phospholipids¹⁵. The acylation of proHlyA does not equate to any of these characterized examples.

The linkage of the radiolabelled fatty-acid group to active haemolysin (HlyA) was resistant to hydroxylamine (not shown), indicating an amide linkage. Acylation of one or more amino groups in proHlyA would explain the increased negative charge of activated HlyA, compared with proHlyA¹⁶. An N-terminal proHlyA acylation similar to *N*-myristoylation seems impossible as there is no specificity for a myristoyl group, nor for the N-terminal sequence as the first amino acid of proHlyA is not

glycine and a mutant of proHlyA lacking the N-terminal 15K is still activated¹⁷. More likely, fatty acylation of proHlyA involves internal amide-linkage (lysine modification), as described for the mouse nicotinic acetylcholine receptor¹⁸, the human insulin receptor¹⁹ and the immunoglobulin heavy-chain polypeptide of mouse B lymphocytes²⁰. Internal proHlyA modification is also consistent with the isolation of a monoclonal antibody which recognizes only the active HlyA, specifically epitope 626-726¹⁷.

Cellular fatty acylation can be central to signal transduction, protein-protein interactions, and protein anchoring to membranes¹². Post-translational modification of proHlyA is not a requirement for HlyB-dependent secretion of HlyA, but it remains possible that it may modulate the kinetics of translocation across the two bacterial membranes. Note that where the exported substrates of eukaryotic homologues of the *E. coli* HlyB translocator are known, they are lipophilic drugs (exported

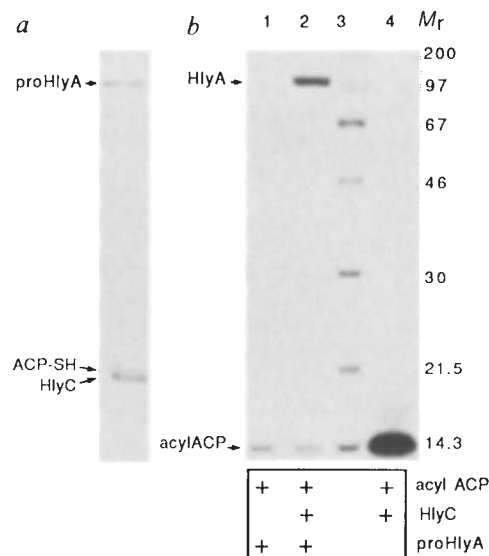


FIG. 2 Activation to the mature toxin occurs concomitantly with the transfer of radiolabelled palmitic acid from ACP to proHlyA. *a*, Coomassie blue-stained SDS-PAGE of reaction mixture components HlyC (2 μg), proHlyA (1 μg) and ACP-SH (3 μg). *b*, Fluorography of *in vitro* radiolabelling reaction mixtures containing the three components, [¹⁴C]palmitoyl-ACP, proHlyA and HlyC, as shown in the box (lanes 1, 2, 4). Lane 3 shows ¹⁴C-labelled *M_r* standards (*M_r* $\times 10^{-3}$).

METHODS. ACP (6 μg , Sigma) was acylated as described in the legend to Table 1 by 1.5 mU acyl-ACP synthetase in 40 μl , for 1 h at 37 °C, in the presence of 100 μM U[¹⁴C]palmitic acid (sodium salt; specific radioactivity 839 mCi mmol⁻¹, Amersham). Then, HlyC was added to 5 μg , proHlyA to 14 μg , and incubations were performed for 30 min at 22 °C in a total volume of 120 μl . Samples (10 μl) were precipitated by 50% (v/v) ethanol (lanes 1 and 2) or 10% TCA (the control, lane 4) and pellets dissolved in SDS-PAGE loading buffer before separation on a 12.5% gel, fluorography and 2 days exposure. Note acyl-ACP migrates faster than ACP-SH, and acyl-ACP is poorly precipitated by 50% ethanol so precipitation was achieved with TCA to maximize recovery in the control lane 4.

by human and parasite multidrug resistance proteins²¹⁻²³) and hydrophobic acylated peptide (yeast a mating factor secreted by the P-glycoprotein STE6 (ref. 24)). What is clear is that proHlyA is not haemolytic^{16,25}, being unable to bind to erythrocyte membranes²⁶. This suggests that ACP-dependent fatty acylation is the key to the targeting of mature HlyA to the mammalian cell membranes, where it may also contribute to the proper folding of HlyA into an active transmembrane pore. ACP-dependent acylation probably activates all the closely related membrane-targeted toxins secreted by pathogenic Gram-negative bacteria, including the adenylate cyclase haemolysin of *B. pertussis*, the causal agent of whooping cough. Toxins of this family require HlyC-type protoxin activation and many have been activated by *E. coli* HlyC^{1,3,27}.

A search for HlyC-like proteins or analogous ACP-dependent protein acylation might reveal broader use of this reaction in other prokaryotic or eukaryotic systems. Whatever its range, the HlyC-dependent activation of proHlyA to the mature membrane-targeted toxin seems to be a novel post-translational modification pathway. We propose that activation is dictated by transfer of a fatty acyl moiety to proHlyA in a process involving association of the negatively charged acyl-ACP with the positively charged HlyC, the latter providing recognition and interaction with specific proHlyA sequences. Kinetic studies *in vitro* have indicated that HlyC, which is produced in the cell in roughly equimolar amounts to proHlyA, can activate only a limited number of proHlyA molecules⁶. Although acyl trans-

ferase activity must presumably reside in HlyC (or in the event of autocatalysis, proHlyA) comparison of sequences with known acyl transferases has revealed no identity. □

Received 9 April; accepted 21 May 1991.

1. Koronakis, V., Cross, M., Senior, B., Koronakis, E. & Hughes, C. *J. Bact.* **169**, 1509-1515 (1987).
2. Glaser, P., Sakamoto, H., Bellalou, J., Ullmann, A. & Danchin, A. *EMBO J.* **7**, 3997-4004 (1988).
3. Gygi, D. et al. *Molec. Microbiol.* **4**, 123-128 (1990).
4. Welch, R. A. *Molec. Microbiol.* **5**, 521-528 (1991).
5. Barry, E. M. et al. *J. Bact.* **173**, 720-726 (1991).
6. Hardie, K. R., Issartel, J.-P., Koronakis, E., Hughes, C. & Koronakis, V. *Molec. Microbiol.* **5** (in the press).
7. Vanaman, T. C., Wakil, S. J. & Hill, R. L. *J. Biol. Chem.* **243**, 6420-6431 (1968).
8. Jackowski, S. & Rock, C. O. *J. Bact.* **169**, 1469-1473 (1987).
9. Therisod, H. & Kennedy, E. P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8235-8238 (1987).
10. Magee, A. I. & Courtneidge, S. A. *EMBO J.* **4**, 1137-1144 (1985).
11. Rock, C. O. & Garwin, J. L. *J. Biol. Chem.* **254**, 7123-7128 (1979).
12. Schmidt, M. F. G. *Biochim. biophys. Acta* **988**, 411-426 (1989).
13. Duronio, R. J., Towler, D. A., Heuckeroth, R. O. & Gordon, J. I. *Science* **243**, 796-800 (1989).
14. Cross, G. A. M. *Cell* **48**, 179-181 (1987).
15. Hayashi, S. & Wu, H. C. *J. Bioenerg. Biomemb.* **22**, 451-471 (1990).
16. Nicaud, J.-M., Mackman, N., Gray, L. & Holland, I. B. *FEBS Lett.* **187**, 339-344 (1985).
17. Pellett, S., Boehm, D. F., Snyder, I. S., Rowe, G. & Welch, R. A. *Infect. Immunity* **58**, 822-827 (1990).
18. Olson, E. N., Glaser, L. & Merlie, J. P. *J. Biol. Chem.* **259**, 5364-5367 (1984).
19. Hedo, J. A., Collier, E. & Watkinson, A. *J. Biol. Chem.* **262**, 954-957 (1987).
20. Pillai, S. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7654-7658 (1987).
21. Gros, P., Croop, J. & Houseman, D. *Cell* **47**, 371-380 (1986).
22. Higgins, C. F. *Nature* **340**, 342 (1989).
23. Hyde, S. C. et al. *Nature* **346**, 362-365 (1990).
24. McGrath, J. P. & Varshavsky, A. *Nature* **340**, 400-404 (1989).
25. Ludwig, A., Jarchau, T., Benz, R. & Goebel, W. *Molec. Gen. Genet.* **214**, 553-561 (1988).
26. Boehm, D. F., Welch, R. A. & Snyder, I. S. *Infect. Immunity* **58**, 1959-1964 (1990).
27. Forestier, C. & Welch, R. A. *Infect. Immunity* **58**, 828-832 (1990).

ACKNOWLEDGEMENTS. We thank E. Koronakis and K. Hardie for their help and discussions during the work. This work was supported by the MRC, Wellcome Trust, AFRC and Royal Society.

Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*

Joseph D. Schrag, Yunge Li, Shan Wu & Mirosław Cygler*

Biotechnology Research Institute, National Research Council of Canada, Montréal, Québec H4P 2R2, Canada

THE Ser-His-Asp triad is a well known structural feature of the serine proteases. It has also been directly observed in the catalytic sites of two lipases, whose high-resolution three-dimensional structures have been determined^{1,2}. Lipases show a wide variety of sizes, substrate and positional specificities, and catalytic rates³. They achieve maximal catalytic rates at oil-water interfaces. The fungus *Geotrichum candidum* produces several different forms of lipases, two of which have been purified to homogeneity^{4,5}. Two lipase genes have been identified, cloned and sequenced^{6,7}. Both code for proteins of 544 amino acids with a total relative molecular mass of about 60,000 (M_r , 60K). The two forms are 86% identical. Their isoelectric points differ slightly, being between 4.3 and 4.6. About 7% of the total M_r is carbohydrate. Until now, only a low resolution structure of GCL has been reported⁸, but no high resolution structure has followed. We now report the three-dimensional structure of a lipase from *G. candidum* (GCL) at 2.2 Å resolution. Unlike the other lipases and serine proteases, the catalytic triad of GCL is Ser-His-Glu, with glutamic acid replacing the usual aspartate. Although the sequence similarity with the other two lipases is limited to the region near the active-site serine, there is some similarity in their three-dimensional structures. The GCL is also an α/β protein with a central mixed β sheet whose topology is similar to that of the N-terminal domain of human pancreatic lipase. As in the other lipases^{1,2}, the catalytic site is buried under surface loops. Sequence comparisons with proteins from the

TABLE 1 Crystallographic data

Derivative	Native	K ₂ PtCl ₄ I	K ₂ PtCl ₄ II	PtEt(NH ₂) ₂	HgCl ₂
No. of measurements	122,564	80,779	81,358	88,115	18,873
No. of unique reflections	25,732	13,313	12,887	13,361	2,296
Completeness	93%	99%	97%	99%	97%
Resolution limit	2.2	2.8	2.8	2.8	5.0
R_{merge}^*	0.035	0.048	0.048	0.036	0.033
$R_{\text{iso}}^{\dagger}$	—	0.15	0.22	0.15	0.31
Phasing power [‡]	—	2.23	1.52	1.07	1.13
$R_{\text{cullis}}^{\S}$	—	0.60	0.71	0.72	0.75
No. of major/minor sites	—	2/3	5/3	1/0	1/1
Mean figure of merit	0.73 for 13,257 reflections	0.57 for 2,630 reflections between 3.5-2.8 Å			

$$* R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$$

$$\dagger R_{\text{iso}} = \sum |I_{\text{deriv}} - I_{\text{native}}| / \sum I_{\text{native}}$$

[‡] Phasing power = r.m.s. F_h/E , where F_h , heavy atom structure factor; E , residual lack of closure.

$$\S R_{\text{cullis}} = \sum |F_{\text{h(obs)}} - F_{\text{h(calc)}}| / \sum F_{\text{h(obs)}}$$

cholinesterase family suggest that they also contain the Ser-His-Glu triad.

Crystals of isoform I of GCL (pI 4.56, as defined in ref. 4) were obtained by the vapour diffusion method using PEG 8000 as a precipitant⁹. The structure was solved by the multiple isomorphous replacement (MIR) method (Table 1) at 2.8 Å resolution and refined to an R -factor of 0.195 at 2.2 Å resolution (see legend to Fig. 1).

The lipase is a single domain molecule (the largest protein domain known) with overall dimensions of about 65 × 60 × 45 Å (Fig. 1). It is an α/β structure with a central 11-stranded mixed β sheet. The sheet strongly twisted and most strands are parallel. A small three-stranded β sheet is nearly perpendicular to the large central sheet. On each side of the large sheet are α helices which run roughly parallel to the β strands. Long loops are found on one side of the β sheet with numerous helices and shorter loops forming a cap along the C-terminal ends of the strands of the large sheet. The structure contains a total of 16 α helices.

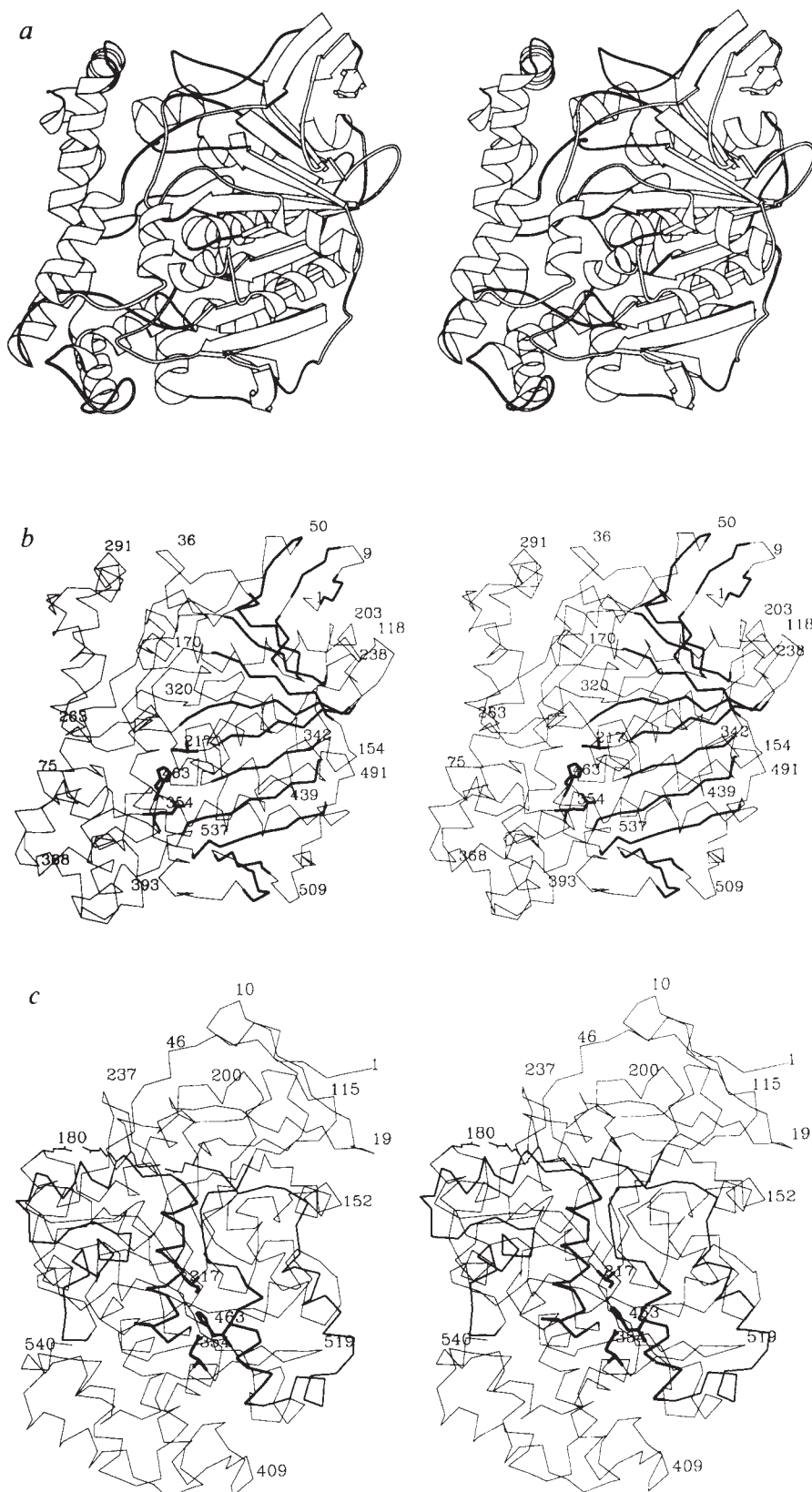
* To whom correspondence should be addressed.

The β -sheet topology of GCL, including the handedness of the connectivities, is very similar to the N-terminal domain of human pancreatic lipase (HPL) and to carboxypeptidase II (ref. 10), another enzyme containing the Ser-His-Asp catalytic triad. The loops and helices connecting many of the strands are, however, quite different. GCL is about 200 residues larger than

the N-terminal region of HPL. The enlarged loops may be involved in interfacial interactions with fatty-acid chains, contributing to substrate specificity.

On the basis of the Gly-X-Ser-X-Gly/Ala consensus sequence common to serine proteases, cholinesterases and lipases¹¹, Ser 217 of GCL has been predicted to be involved in the cataly-

FIG. 1 Stereo view of the tracing of the polypeptide chain. The structure was determined from the monoclinic crystal form ($P2_1$) with cell dimensions $a=59.4$, $b=84.0$, $c=56.0$ Å, $\beta=100.1^\circ$, one molecule per asymmetric unit. Four heavy-atom derivatives have been used to calculate MIR phases (Table 1). Anomalous data proved to be very useful. The mean figure of merit to 2.8 Å resolution was 0.73. These phases were subsequently improved by solvent flattening²⁴. The resulting map was of good quality and showed a number of α helices and β strands. About 80% of the polypeptide chain was traced in this map and modelled as alanines with the help of programs O²⁵ and FRODO²⁶. This partial structure was subjected to one cycle of molecular dynamics simulated annealing refinement using MIR phases as restraints (X-PLOR²⁷). The R -factor dropped from 45% to 35% with data to 2.8 Å resolution. Phases obtained from this model were combined with the MIR phases and the model was refitted to the new map. A unique amino-acid sequence was identified in one region of the map and allowed for correct sequence assignment through the entire length of the chain. In the places where the two gene sequences differed, the electron density agreed consistently with one of the lipase genes⁷. Subsequent molecular-dynamics refinement, which included 8–2.2 Å resolution data, reduced the R factor to 0.245. The calculated phases were again combined with the MIR phases and subsequent refitting was done using the phase combined, F_o-F_c and $3F_o-2F_c$ maps. The current R factor is 0.195 for 23,289 reflections with $I > 2\sigma(I)$. The r.m.s. deviation in bond lengths and bond angles from their ideal values are 0.015 Å and 3.0° , respectively, and the average B factor is 12.2 Å^2 . **a**, Drawing of the protein showing a side view of the 11-stranded mixed β sheet. **b**, α tracing (view as above) with the residues of the catalytic triad shown in full. **c**, View showing the two α helices (heavy lines) covering from the top the putative active site.



sis^{6,7}. This is consistent with chemical labelling and modification studies of pancreatic lipase¹²⁻¹⁴, cutinase^{15,16} and cholinesterases¹⁷⁻²⁰, which conclusively demonstrated that Ser in the consensus sequence is the catalytic residue in these enzymes. Chemical modification studies on pancreatic lipase further indicated the requirement of a histidine and a carboxyl group²¹. This was confirmed by the presence of a Ser-His-Asp triad in the structure of HPL¹. In the three-dimensional structures of the two lipases^{1,2} and carboxypeptidase II¹⁰ the catalytic serine is located at a tight turn between the C terminus of a β strand and the N terminus of an α helix. Ser 217 in GCL is located in an identical topology. His 463 and the carboxyl group of Glu 354 complete this putative catalytic triad. Although one would expect the carboxyl moieties of glutamate and aspartate to be

chemically equivalent, this is the first observation of a glutamate rather than an aspartate in this role.

Comparison of GCL, HPL, *Rhizomucor miehei* lipase and carboxypeptidase II shows that in all of them the sequential order of active site residues is Ser, Asp/Glu, His. Although the serine maintains its location, the histidine and the carboxylate come from different loops in these proteins.

Superposition of the catalytic triads of GCL, subtilisin, *Streptomyces griseus* trypsin, chymotrypsin and papain (Fig. 2) indicates that in GCL, as in papain, the nucleophile is on the opposite side of the histidine ring than in serine proteases²². The oxyanion hole in GCL is most probably formed by the backbone nitrogens of Ala 218 and Ala 132.

As in the other two lipases, the catalytic site in GCL is

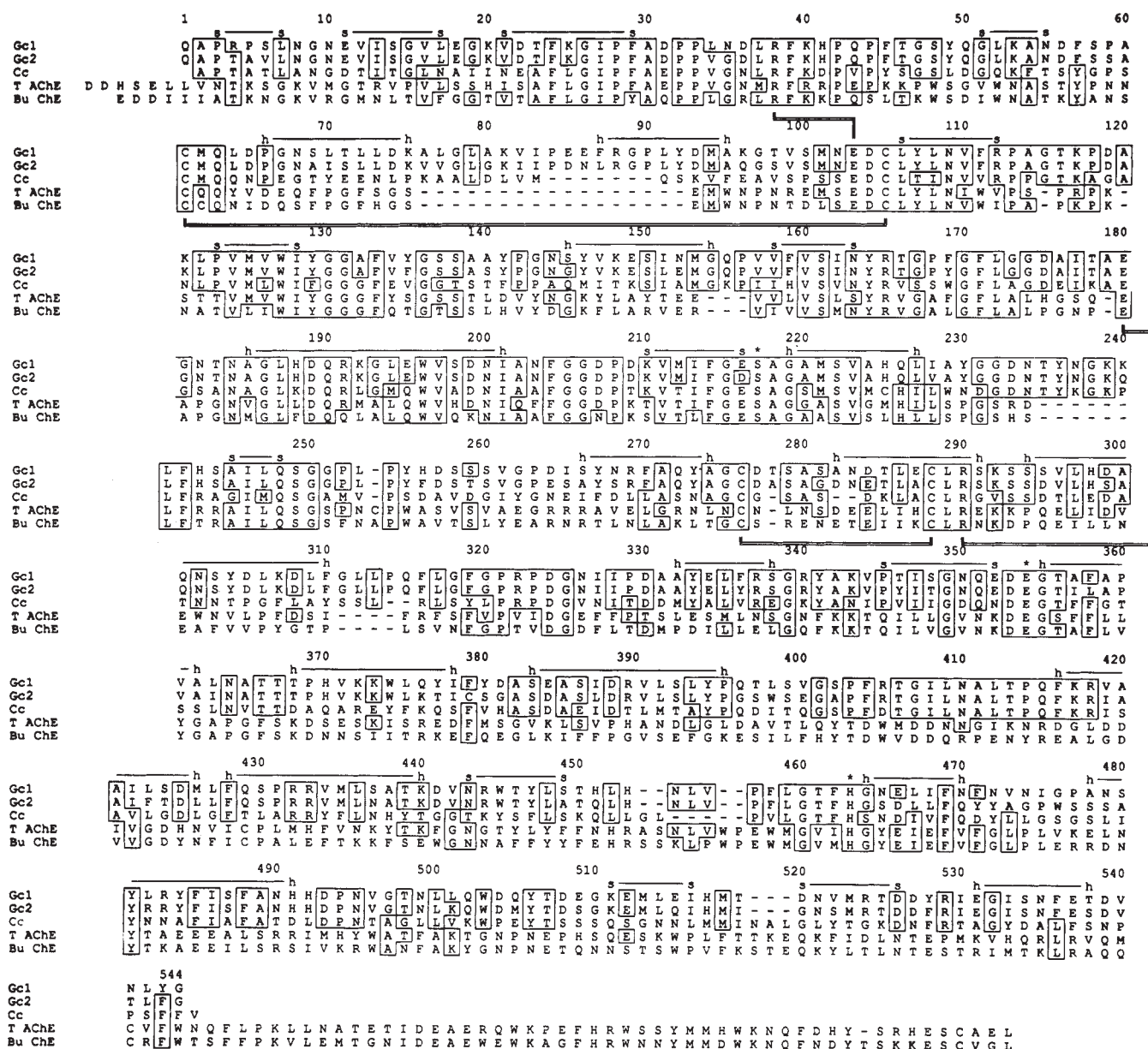


FIG. 3 Sequence alignment of GCL isoform I⁷ (Gc1), isoform II⁶ (Gc2), *C. cylindracea* lipase²³ (Cc), *Torpedo californica* acetylcholinesterase¹⁸ (T AChE) and human butyrylcholinesterase¹⁹ (Bu ChE). Residues that are identical in at least three sequences are boxed (single-letter amino-acid

notation). The secondary structure elements observed in GCL are shown above the sequence. α helix is shown as h—h; β strand is shown as s—s; *, residues from the catalytic triad; the pattern of Cys-Cys bridges and conserved salt bridges is also shown.

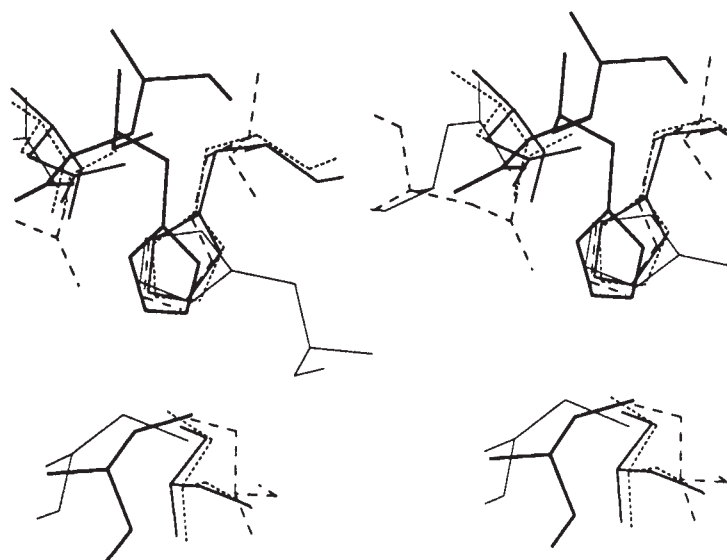


FIG. 2 Stereo view of the superposition of triad from GCL (thick lines) with *S. griseus* trypsin (medium lines, entry 1SGT in the Brookhaven Data Bank²⁸), papain (thin lines, entry 6PAD in Protein Data Base (PDB)), subtilisin (long dashed lines, entry 2SEC in PDB) and chymotrypsin (short dashed lines, entry 6CHA in PDB).

inaccessible from the surface. Ser 217 is buried even deeper in the protein than the equivalent serines in the other lipases. This is possibly related to this enzyme's preference for long-chain fatty acids. The putative active site is covered by two nearly parallel α helices (residues 66–76 and 294–310) coming from two different surface loops (Fig. 1c). The outside surfaces of these helices and a bordering region are relatively hydrophobic, suggesting that this might be the surface through which GCL binds to the interface. If so, one or both of these helices must move to allow substrate to gain access to the catalytic site. Helix 66–76 is the first in a helix-loop-helix twisted flap. This twist at the base of the flap is stabilized by a disulphide bond and a salt bridge. It is tempting to suggest that a simple unwinding of the twist would open a channel to the catalytic site. That, however, would move the second helix away from the core of the protein, and helix 66–76 would still restrict access to Ser 217. The other helix, 294–310, covering the active site is contained in a larger loop (residues 262–311) which is bordered by glycine residues. Although these glycines may provide some flexibility and allow a shift on contact with an interface, this loop also has a salt bridge between Arg 290 and Glu 180. Perhaps access to the catalytic serine is achieved by small shifts of a few structural elements in this region rather than a major shift of a single feature.

There is also a second triad arrangement comprised of Ser 449–His 451–Asn 465. This triad is also buried. Unlike catalytic triads in other enzymes, all three residues are very close in the amino-acid sequence and Ser 449 is located near the end of a β strand rather than at a strand-helix junction. Sequence comparison with related proteins (Fig. 3) shows that this triad is not conserved. As there is no biochemical evidence for a second active

site, the formation of this second triad may just be coincidental. Such fortuitous triads have been found in other proteins²³.

GCL seems to be very hydrated. Many water molecules are also observed in the interior of the molecule. Of special interest is the presence of two acidic residues, Glu 216 and Glu 466, near the catalytic site. They coordinate two water molecules close to the gamma oxygen of Ser 217. These residues are either Glu or Asp in other lipases and in cholinesterases. We speculate that their role is to provide a trap for water molecules which are essential for hydrolysis.

Sequence comparisons have shown that the *G. candidum* lipases are related to the lipase from *Candida cylindracea* and to the cholinesterases⁷. Alignment of these sequences (Fig. 3) indicates that many structural features are conserved. The sequence identities are very apparent in the β -strand regions of the GCL structure. The *C. cylindracea* lipase and the acetylcholinesterases probably have a central β sheet which is nearly identical to that seen in GCL. The disulphide bonding pattern, the charged and polar groups in the region of the catalytic site, the salt bridges which stabilize the surface loop between residues 61 and 105, and most importantly, the Ser-His-Glu triad are all conserved. The catalytic triad in *C. cylindracea* is probably Ser 209–His 449–Glu 341 and for the *Torpedo* acetylcholinesterase Ser 200–His 440–Glu 327. The major differences between the lipase and cholinesterase sequences all occur in regions corresponding to loops in the GCL structure and many are differences in the lengths of the loops. The cholinesterase sequence homology is weaker in the C-terminal third of the sequence. This is not too surprising as acetylcholinesterases form covalent dimers through a disulphide bridge involving a cysteine residue near the C terminus²⁰. □

Received 22 March; accepted 16 May 1991.

- Winkler, F. K., D'Arcy, A. & Hunziker, W. *Nature* **343**, 771–774 (1990).
- Brady, L. et al. *Nature* **343**, 767–770 (1990).
- Borgstrom, B. & Brockman, H. L. *Lipases* (Elsevier, Amsterdam, 1984).
- Veeraragavan, K., Colpitts, T. & Gibbs, B. F. *Biochim. biophys. Acta* **1044**, 26–33 (1990).
- Sugihara, A., Shimada, Y. & Tominaga, Y. *J. Biochem.* **107**, 426–430 (1990).
- Shimada, Y., Sugihara, A., Tominaga, Y., Iizumi, T. & Tsunasawa, S. *J. Biochem.* **106**, 383–388 (1989).
- Shimada, Y., Sugihara, A., Iizumi, T. & Tominaga, Y. *J. Biochem.* **107**, 703–707 (1990).
- Hata, Y. et al. *J. Biochem.* **86**, 1821–1827 (1979).
- Schrag, J. D., Li, Y., Wu, S. & Cygler, M. *J. molec. Biol.* (in the press).
- Liao, D.-I. & Remington, S. J. *J. biol. Chem.* **265**, 6528–6531 (1990).
- Brenner, S. *Nature* **34**, 528–530 (1988).
- Chapus, C., Sémériva, M., Bovier-Lapierre, C. & Desnuelle, P. *Biochemistry* **15**, 4980–4987 (1976).
- Garner, C. W. *J. biol. Chem.* **255**, 5064–5068 (1980).
- Hadvary, P., Sidler, W., Meister, W., Vetter, W. & Wolfer, H. *J. biol. Chem.* **266**, 2021–2027 (1991).
- Koller, W. & Kolattukudy, P. E. *Biochemistry* **21**, 3083–3090 (1982).

- Ettinger, W. F., Thukral, S. K. & Kolattukudy, P. E. *Biochemistry* **26**, 7883–7886 (1987).
- Yamato, K. et al. *Biochem. Genet.* **21**, 135–145 (1983).
- MacPhee-Quigley, K., Taylor, P. & Taylor, S. *J. biol. Chem.* **260**, 12185–12189 (1985).
- Lockridge, O. et al. *J. biol. Chem.* **262**, 549–557 (1987).
- Sikorav, J.-L., Kejci, E. & Massoulié, J. *EMBO J.* **6**, 1865–1873 (1987).
- Chapus, C. & Sémériva, M. *Biochemistry* **15**, 4988–4991 (1976).
- Garavito, R. M., Rossmann, M. G., Argos, P. & Eventoff, W. *Biochemistry* **6**, 5065–5071 (1976).
- Blow, D. L. *Nature* **343**, 694–695 (1990).
- Wang, B. C. in *Meth. Enzym.* **115**, 90–112 (1985).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1991).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 614–617 (1978).
- Brunker, A. T., Karplus, M. & Petsko, G. *Acta crystallogr.* **A45**, 50–61 (1989).
- Bernstein, F. C. et al. *J. molec. Biol.* **112**, 535–543 (1977).
- Kawaguchi, Y., Honda, H., Taniguchi-Morimura, J. & Iwasaki, S. *Nature* **341**, 164–166 (1989).

ACKNOWLEDGEMENTS. We thank M. Desrochers for technical help, E. Brändli for assistance in lipase purification and A. Storer for helpful discussions. The Ca coordinates and the coordinates of catalytic residues will be deposited in the Brookhaven Databank.

Nanovid microscopy

Hugo Geerts, Mark de Brabander and Rony Nuydens

By combining small colloidal gold probes with video-enhanced quantitative microscopy, the intracellular dynamics of specific proteins in living cells can now be studied.

EVER since Anton Van Leeuwenhoek opened up the microscopic world to science, researchers have tried to probe deeper and deeper into the secrets of nature by increasing the resolution of light microscopy. Ultimately, the physical limit of resolution was calculated to be in the order of the wavelength of light. For many years, electron microscopy, in combination with immunogold techniques, provided a resolution down to the level of the single molecule. But, the intracellular dynamics of proteins within living cells seemed virtually out of reach.

Video-enhanced microscopy

In 1981, Allen¹ and Inoué² independently discovered the concept of video-enhanced microscopy. Structures down to 20 nm were visualized on a standard research microscope by use of a video camera. The idea was to open the condenser aperture to increase the resolution. The dramatic fall in contrast that resulted, caused by the high light intensity, was reverted by a two-step procedure of electronic contrast enhancement. First, a d.c. signal, called an offset, was subtracted from the analog video signal and the remaining difference was electronically amplified. However, electronic amplification also enhanced imperfections in the illumination profile or dust particles in the optic pathway. As a final step, a 'mottle' image, acquired with the sample out of focus, was subtracted digitally from the analog-enhanced video image.

In this way, minute structures down to 20 nm, such as microtubules or vesicles in living cells, became clearly visible. However, this was not a violation to Abbe's law. Although the minute structures in the image plane were only 20–30 nm in size, their image on the monitor screen appeared 'blown up' by the point spread function of the optical system to a size equivalent to the image of a point (300–400 nm). This means that if two such structures were less than a wavelength apart, they merged into one structure and were indistinguishable as independent entities. However, in the case of low-density gold labelling, it was actually possible to see individual 20-nm gold probes.

Two limitations still remained. First, researchers wanted to quantify dynamic changes in the intracellular pattern of vesicles and structures. With Normarski interference or phase-contrast microscopy, this was not an easy task for a computer. As Normarski interference or phase-contrast images are far more complex than simple brightfield images of black dots on a white background, it is much more difficult to

detect motion. However, humans can easily detect motion in such sequences of images. The algorithms used by humans in the detection of motion have not been elucidated thus far. In the case of nanovid microscopy, the computer can accurately calculate the positions of the colloidal gold particles in a time series of video images, thereby providing information on the motion. In a few specific cases, such as the fast axonal transport of vesicles in neuronal cell cultures, computer algorithms for automatic quantification have been developed³.

The second limitation concerned specificity. The development of confocal and quantitative microscopy with fluorescently-labelled antibodies or fluorescent environment-sensitive probes, that is probes like fura-2 or BCECF, has led to the dynamic localization of very well-defined molecules, from ions to proteins. Here, however, spatial and time resolution are not compatible. Confocal microscopy provides superior spatial resolution in three dimensions, but at the expense of time resolution. And, sensitive photodiodes easily resolve dynamic changes in the millisecond range, but without any spatial information. Furthermore, photobleaching and subsequent phototoxicity severely limit their use in the study of dynamic phenomena within the cell.

Nanovid microscopy

The use of small colloidal gold probes of 20–40-nm diameter in combination with video-enhanced microscopy provides a new research tool for following intracellular dynamics in living cells⁴. Indeed, colloidal gold probes down to 20 nm can be visualized using a video-enhanced brightfield microscope because they scatter enough light out of the aperture of the objective (see Fig. 1). When coupled to specific antibodies, these gold probes labelled specific proteins. Because of their punctuate appearance, it is now possible to follow the intracellular dynamics of specific molecules, which can be precisely localized with regard to morphological features such as cell boundary or cell nucleus. In accordance with Abbe's law, when two or more gold particles are closer than a wavelength they merge into one cluster with increased contrast. As most of the light energy passes through the sample without any absorption, little phototoxicity is observed. Consequently, dynamic phenomena can be followed in real-time.

Applications of the technology

In vitro mobility assays. Nanovid microscopy has been used to study the mobility of

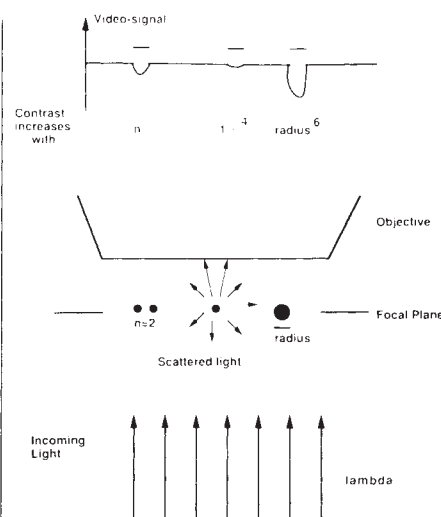


FIG. 1 Physical principle of image formation in nanovid microscopy. In transmission mode, the light is scattered by the colloidal gold particle out of the aperture of the objective. This small difference in intensity is further enhanced so that the gold appears as small black dots against an essentially white background. The intensity, or 'blackness' of a cluster, increases with the number and size of the particles.

40-nm, BSA-PEG-stabilized colloidal gold probes powered by an extract containing kinesin on polymerized microtubules⁵ (see Fig. 2a). This provided an elegant approach to studying the effect of certain interventions on kinesin-mediated 'vesicle' transport in a reconstituted system.

Receptor-mediated endocytosis. The nanovid technique has been used to study the fate of 25-nm colloidal gold probes coupled to a monoclonal antibody against the transferrin receptor in A 431 cells⁶ (see Fig. 2b). This is believed to be the first time that receptor-mediated endocytosis has been followed in real-time.

Motion of cell-surface molecules. The motion of discrete 40-nm colloidal gold particles coupled to poly-L-lysine or to monoclonal antibodies against glycoproteins on the surface of cells has been studied^{7,8} (see Fig. 2c). The gold coupled to an antibody against the lipid-linked Thy-1 antigen or to short poly-L-lysine molecules exhibited a prominent Brownian motion, in accordance with data obtained by video-FRAP (fluorescence recovery after photobleaching) measurements. In addition, corralled diffusion in microdomains was observed. Conversely, gold probes coupled to anti-pgp-80 glycoprotein or to long poly-L-lysine molecules were actively driven rearward by a microfilament network.

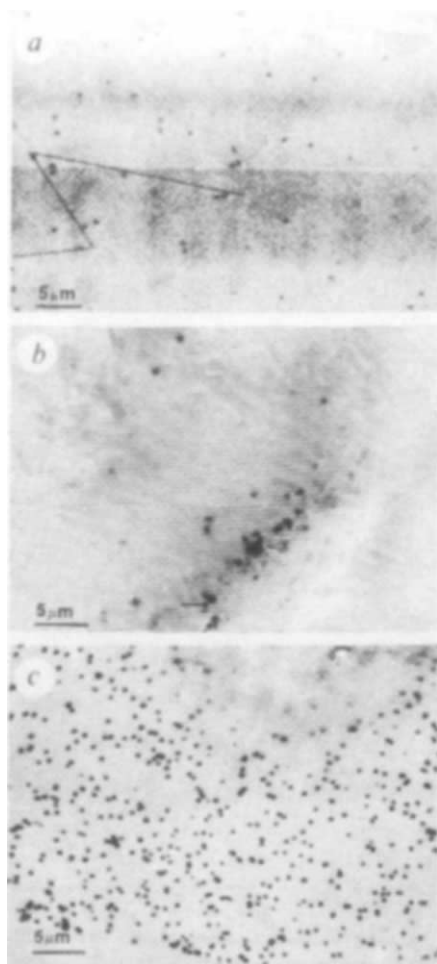


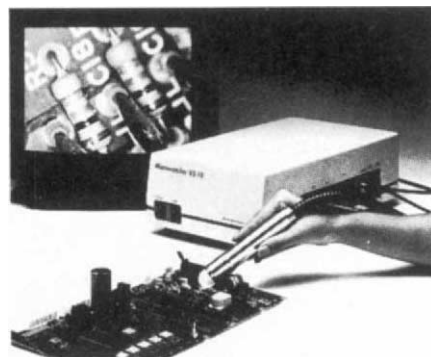
FIG. 2 Applications of nanovid microscopy. *a*, Mobility of gold probes on polymerized, taxol-stabilized microtubules powered by kinesin. The pathlength traversed by the particle within 120 seconds is presented. *b*, Binding and uptake of anti-transferrin coupled to 25-nm gold on A431 cells. Particles bind immediately to the peripheral rim of the cell surface and, within a few minutes, some of them are taken up by endocytosis, which is evident by their typical saltatory motion. The large vesicle (see large arrow) arises from fusion of several smaller vesicles and contains gold particles bound to the membrane of the organelle. *c*, Distribution of 4 kD poly-L-lysine coupled 40-nm colloidal gold particles on the surface of a PTK₂ cell. Notice the distinctive punctuate labelling, which opens up new possibilities for the quantitative localization of cell-surface proteins.

Axonal transport. Nanovid microscopy has also been used to measure axonal transport in neuronal cells. It has been shown that 30-nm, BSA-PEG-stabilized gold particles are actively taken up by neuroblastoma cells and move with a typical saltatory motion, which is characterized by long linear jumps interleaved with longer stop times (see Fig. 3). More recently, nanovid technology has been used to study 30-nm gold probes coupled to a monoclonal antibody against the nerve growth factor receptor on neurons. This opens up considerable possibilities for the study of axonal transport in various models of neuropathy. □

Picture perfect

Focusing this week on microscopy and image analysis, new products include a hand-held microscope, a new objective that converts a standard microscope to a fluorescence microscope, and a DNA sequence gel reader with voice recognition.

THE Mitsubishi Kasei Microwatcher Series of **hand-held microscope systems** can be used for applications as diverse as microelectronics to biotechnology (*Reader Service No. 101*). Two microscope systems in the Microwatcher Series are available, the VS-10 and the VS-30, both of which are compact, lightweight and portable, and integrate precision



Mitsubishi Kasei's hand-held microscope.

fibre optics with video technology. To use, the handy CCD sensor with its polarized lens is passed over the specimen. Images are viewed on a CRT monitor. Seven interchangeable optical polarized lenses ($\times 10$, $\times 20$, $\times 50$, $\times 100$, $\times 200$, $\times 500$ and $\times 1,000$) provide the Microwatcher with a wide magnification range.

Add-on accessories

Carl Zeiss has added a new 35-mm camera, the MC 80, to its line of **microscope cameras** for automatic photomicrography (*Reader Service No. 102*). Although compatible with all Zeiss microscope systems, the camera is particularly suited for use with the Zeiss Standard 25 routine microscopes, the Axioskop 20 and the stereomicroscopes. The MC 80 features a vibration-free shutter, 100 per cent light-to-film measurement, and a $\times 2.5$ projection lens for direct, high-contrast imaging onto film. Additional design features include auto exposure with 35 per cent centre-weighted metering, and motorised film advance and rewind. Exposure time ranges from 1/125th to 250 seconds in auto mode, extending to 4,000 seconds in manual mode. For specific applications, the exposure time can be stored or manually set; the film can also be double exposed, which is useful for double fluorescence. The camera accepts film ranging from 25 to 6,400 ASA. An optional data back with a 5-mm reference scale permits the date and time to be recorded on the film.

Polaroid's new TX-1000 **thermal image printer** incorporates the latest advances in thermal dye diffusion transfer technology to produce A6 (14 $\times$ 10-cm) colour prints or

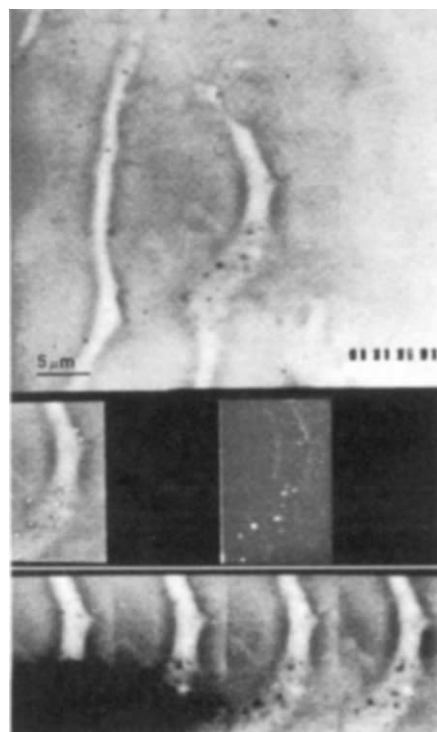


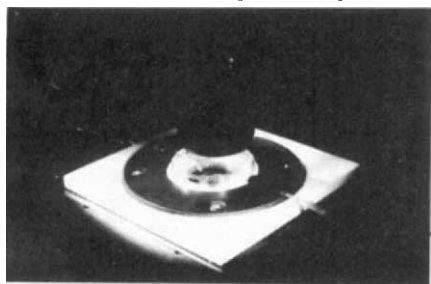
FIG. 3 Colloidal gold particles (30 nm) taken up within vesicles by N4 neuroblastoma cells can be used to study fast axonal transport. Some of the steps involved in the automatic quantification of colloidal gold are shown here. From the complete image (top), a small area of interest is copied successively in different parts of the image processor (bottom). Each area in the image is then analysed to detect the individual gold particles (middle).

Hugo Geerts, Mark de Brabander and Rony Nuydens are in the Department of Physiology, Life Sciences, Janssen Research Foundation, Beerse, Belgium. For more information, fill in reader service number 100.

1. Allen, R., Travis, J., Allen, N. & Yilmaz, H. *Cell Motil. Cytoskel.* **1**, 275-289 (1981).
2. Inoué, S. *J. Cell Biol.* **89**, 346-356 (1981).
3. Cornelissen, F., Nuydens, R., Nuydens, R. & Geerts, H. *J. Neurosci. Meth.* **35**, 79-88 (1990).
4. De Brabander, M., Nuydens, R., Geuens, G., Moeremans, M. & DeMey, J. *Cell Motil. Cytoskel.* **6**, 105-113 (1986).
5. Geerts, H. *et al. Biophys. J.* **52**, 775-782 (1987).
6. De Brabander, M., Nuydens, R., Geerts, H. & Hopkins, C. *Cell Motil. Cytoskel.* **9**, 30-47 (1988).
7. De Brabander, M. *et al. J. Cell Biol.* **112**, 111-124 (1991).
8. Sheetz, M.P., Turney, S., Qian, H. & Elson, E.L. *Nature* **340**, 284-288 (1989).

transparencies with a resolution of 138 dpi and 262,000 colours (*Reader Service No. 103*). These features make the TX-1000 suitable for applications requiring high resolution and subtle colour tonality, says Polaroid. In addition to its tilting and mirror image capabilities, the TX-1000 also features built-in frame storage, reverse imaging, a multiple print capability, and wireless remote control. Standard TV controls allow the operator to lighten or darken the final print, and adjust contrast or colour intensity. Alternatively, an infra-red remote control permits operation of the TX-1000 from a distance. The system is compatible with PAL video, super VHS and RGB signals. Polaroid believes that the thermal printer will have broad application in scientific and medical fields. Prices for the TX-1000, which will be available in July, start at £1,350 (UK).

The new **microincubation chamber** from BTI provides viewing, access and thermoregulation to tissues or cultured cells on inverted microscopes (*Reader Service No. 104*). The system supports thermoregulation from 4–45 °C with ± 0.2 °C precision, in addition to micromanipulation, perfusion,

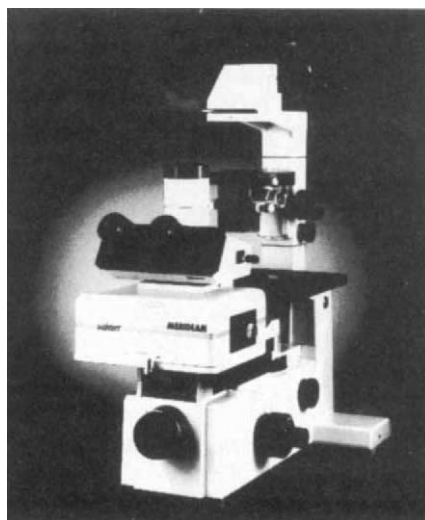


BTI's microincubation chamber supports temperature control from 4–45°C.

atmospheric control, and long-term microscopical studies under sterile conditions. The ultrathin (0.07 mm) coverslip well bottom enhances the use of short working-distance objectives on inverted microscopes. The coverslips also fit in standard 35-mm culture dishes. The BTI chamber is also suitable for use with popular fluorescent optical indicators such as fura-2, indo-1 and BCECF for measuring intracellular calcium and pH.

Fluorescence microscopy

INSIGHT is the latest **laser scanning confocal microscope** from Meridian Instruments, which is designed to provide high-resolution images at video rates when viewed directly or via a CCD camera (*Reader Service No. 105*). The INSIGHT system is supplied with a scanner, research-grade inverted microscope, air-cooled laser and 35-mm camera. The system uses patented Bilateral Scanning, a new technology using a unique double-sided, rapid-scanning mirror for viewing confocal sections of living cells in real-time, says Meridian. It works as follows: as a narrow cursor of light illuminates the field of view, a double-sided, galvanometer-driven mirror simultaneously scans the specimen and detector (that is, the oculars or video



Video rate confocal imaging.

camera) at rates of up to 120 frames per second, building a confocal image. The rapid scanning speed permits direct ocular viewing of the phase-contrast and fluorescent confocal images, individually, or superimposed as phase-fluorescent image overlays, says Meridian. In addition, phase and fluorescent confocal images can be viewed via an optional video camera. INSIGHT sells for under \$55,000 (US).

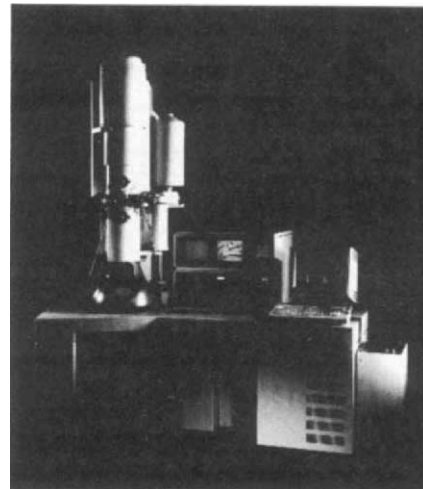
K.W. Kirk, in collaboration with Flow, has developed a new type of **microscope objective**, the Mackler objective, which provides an inexpensive way of upgrading light microscopes to fluorescence microscopes (*Reader Service No. 106*). The Mackler objective, named after its inventor, Dr M.T. Mackler, has integral barrier beam splitter and excitation filters and is simply screwed into the microscope objective turret like any standard objective, says Kirk. Four sizes of objective are available: $\times 10$, $\times 40$, $\times 60$ and $\times 100$ (oil immersion). Illumination is provided by a flexible guide and a 150-W light source.

SEMs/TEMs and EC-AFMs

The JSM-5400 is the new general purpose, **digital scanning electron microscope** (SEM) from Jeol, which is designed for operation by inexperienced and experienced users alike (*Reader Service No. 107*). The SEM features high resolution to 4 nm and a magnification range from $\times 15$ to $\times 200,000$. The SEM has been designed with ease of operation in mind; operations from vacuum pumping to image observations and photography are fully automated. The instrument's framestore function allows even grainy pictures, such as high resolution and non-evaporation low-voltage images, to be converted into bright and easy-to-view still pictures, Jeol says. Behind all these features, which are designed to simplify operation, lies the instrument's new electron optical system. According to Jeol, the zoom condenser lens causes very little field shift, defocusing, and no change in image brightness or contrast even when the probe current is changed. In

addition, the variable objective lens aperture is heated constantly to keep it contamination-free. For analytical studies, the JSM-5400 can be fitted with both energy-dispersive or wavelength-dispersive X-ray spectrometers.

The CM20 FEG is the new **transmission electron microscope** (TEM) from Philips Electronics that has a nano-analysis capability (*Reader Service No. 108*). The instrument is designed around a field emission source with a brightness that is 1,000 times greater than a conventional filament, Philips says. Increased angular and spatial resolu-



Nano-analysis in the TEM mode.

tion in techniques ranging from convergent beam diffraction to energy-dispersive X-ray spectrometry and EELS analysis are combined with high-resolution imaging to provide a fully integrated system for microscopy on the nanometre scale. The user interface of the CM20 FEG provides full and easy control of the field emission source and also continuously monitors all critical parameters and the automated start-up/shut-down procedures. Philips states that experienced operators for the TEM are not necessary. The flexible hardware and upgradable software allow the CM20 FEG to be tailored to suit specific applications in fields as diverse as semiconductor technology, polymer structure research, metallurgy, macromolecular science and the development of new ceramic and superconducting materials.

Using the NanoScope **electrochemical atomic force microscope** (EC-AFM) from Digital Instruments, it is possible to image surfaces undergoing electrochemical reactions with an electrically neutral probe (*Reader Service No. 109*). The EC-AFM combines Digital's AFM with a potentiostat-controlled electrochemical cell, allowing simultaneous AFM imaging and voltammetry with atomic resolution. The instrument controls are straightforward and fully interactive; microscope and potentiostat controls can be changed without interrupting the experiment, says Digital. The EC-AFM has been used to observe the deposition of

single atomic layers, imaging of adsorbates, surface diffusion, nucleation, and corrosion processes. Digital supplies the EC-AFM as an attachment to the NanoScope system.

Image processing

The DNA Parrot talking gel reader from T & T Research now features voice recognition (*Reader Service No. 110*). Users can now read a sequence into a microphone instead of pushing buttons on a cursor for data entry into a Macintosh or IBM PC or compatible computer. The gel reader consists of a serial port cursor device with four buttons for sequence data entry and one for deletion of the last entry. Sequence data is entered by moving the cursor upward on the

gel and pressing the respective buttons, or, in the case of the voice recognition option, reading the sequence into a microphone. As the button is pressed or the base pronounced, it is recognised and entered onto the screen. To confirm an entry, the system provides both audio feedback in the form of digitized sound and visual feedback by the appropriate button on the cursor lighting up. The visual feedback signal is provided irrespective of whether the entry was made by pressing the button or by voice recognition. A complete DNA degenerate alphabet is supported and the data can be exported as a plain text or an IntelliGenetics file.

A newly developed densitometer software package called Gelscan from Günter Spieker, when used in conjunction with a commercially available scanner, can be used to evaluate PAGE-gels (*Reader Service No. 111*). The set up, which can also be used to evaluate immunoblots and autoradiograms, uses the GT-6000 scanner from Epson and costs DM14,500 (Germany). The scanner works in the reflection mode and so will also accept photographic prints and other opaque originals. Three scan ranges are provided for optimizing sensitivity and linearity. Four scan colours are available and can be selected according to the dyes used. The resolution is 600 dpi, says the company. The program, which runs under Windows 3.0, can be used, for example, for peak detection, molecular weight determination, peak area integration and for the overlay of up to eight scans.

Cream is a new PC-based software package from Bio Trek that has been designed for the **automatic quantification of electrophoretic patterns** (*Reader Service No. 112*). Cream digitizes images by a video camera and a framegrabber at a resolution of 512 × 512 pixels in 256 grey levels. The images are displayed in 640 × 480 pixels with 256 colours on a VGA system. The digitization time is less than 0.1 second, says Bio Trek. Operation of the Cream program is designed to be simple: a suitable sub-menu is selected, the object to be counted is placed under the video camera and, at the push of a button, an image of the object is 'grabbed' by the computer and displayed on the screen. At this point the operator is able to work with the image, that is, amplify discrete regions, and quantitate the size and intensity of objects. Images can be stored on disk, exported in TIFF to wordprocessors or printed in up to 64 grey levels on a laser printer. Numerous application modules are available, including modules for scanning one-dimensional gels (SDS-gels, western blots, autoradiograms) and two-dimensional gels, reading ELISA plates, quantifying CIE-gels (crossed immunoelectrophoresis images), evaluating Mancini rings, counting bacterial colonies and reading DNA sequences.

Version 1.4 of the Java **video analysis soft-**

ware package for IBM PCs and compatibles has just been launched by Jandel Scientific (*Reader Service No. 113*). Like its predecessor, Version 1.4 includes image processing, densitometry, morphometric analysis and data analysis, in addition to new features designed to enhance the overall program execution, including image arithmetic, image snapshots, image bit map transforms, enhanced macro facilities, and data worksheet formatting and printing. Version 1.4



Video view — see Jandel.

also uses available expanded memory to free conventional memory, and enhance the program's new features like snapshots and image arithmetic. One of the new features, image bit map transforms, allows image pixel intensity modifications to be transferred to and from the data worksheet and frame grabber buffer. When Version 1.4 is running with expanded memory, up to nine images can be saved as snapshots for quick access. The new image arithmetic feature lets the user add, subtract and average images. If images are saved to disk, they can be added, subtracted and averaged with the current image in the frame grabber buffer. Image arithmetic also measures changes over time and can improve the signal-to-noise ratio of an image. Version 1.4 costs \$1,495 (US), although owners of Version 1.3 can upgrade to Version 1.4 for \$295 (US).

High throughput and greater accuracy are the selling points of the new **colony counting system** from SeeScan, the company that developed the product (*Reader Service No. 114*). The system is designed to provide fast counts of pour and spiral plates and can be used to count difficult colonies including pin-points, spreaders and touching colonies. The counter can process up to 600 plates per hour, compared to an average of 60 plates per hour with manual counting methods, says the company. The colony counting system consists of the counter, monitor and keyboard. The counter is operated from the keyboard by a menu-driven system that includes on-screen help messages to guide the operator. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

Synthetic Oligonucleotides

Custom made probes and primers, synthesised, purified and shipped in 3 days

£3.50 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

Modified bases, biotin or digoxigenin labelling, data base analysis and full design facilities

Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX

Tel: XX-44-(0)71 - 326 3126

FAX XX-44-(0)71 - 733 3877



The original
**SINGER MK I
MICROMANIPULATOR**
single or in pairs

FOR

- XENOPUS OOCYTE INJECTION
- PIPETTING OF CELLS IN SUSPENSION
- MICRODISSECTION AND MICRO SURGERY

INT. SALES: +44.984.40226

FAX: +44.984.41166 for details

Custom Peptides

\$30 a residue, 10 µ mole scale
\$80 a residue, 100 µ mole scale

Research Genetics

1-800-533-4363

In the United Kingdom 0 800 89 1393

In Canada 1 800 533 4363

FAX 205-536-9016